

urdukutabkhanapk.blogspot

21st century
first
REFERENCE SERIES

21st Century POLITICAL SCIENCE

A Reference Handbook



Edited by

John T. Ishiyama
Marijke Breuning

21st Century
POLITICAL
SCIENCE
A Reference Handbook

21st Century POLITICAL SCIENCE *A Reference Handbook*

Edited by

John T. Ishiyama
Marijke Breuning

University of North Texas



reference

Los Angeles | London | New Delhi
Singapore | Washington DC

Copyright © 2011 by SAGE Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

For information:



SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London, EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/1 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia Pacific Pte. Ltd.
33 Pekin Street #02 01
Far East Square
Singapore 048763

Printed in the United States of America

Library of Congress Cataloging in Publication Data

21st century political science : a reference handbook / editors, John T. Ishiyama, Marijke Breuning.
p. cm.

Includes bibliographical references and index.

ISBN 978 1 4129 6901 7 (cloth)

1. Political science Handbooks, manuals, etc. I. Ishiyama, John T., 1960 II. Breuning, Marijke, 1957

JA66.T94 2011

320 dc22 2010020638

This book is printed on acid free paper.

10 11 12 13 14 10 9 8 7 6 5 4 3 2 1

<i>Publisher:</i>	Rolf A. Janke
<i>Assistant to the Publisher:</i>	Michele Thompson
<i>Acquisitions Editor:</i>	Jim Brace Thompson
<i>Developmental Editor:</i>	Sanford Robinson
<i>Reference Systems Manager:</i>	Leticia Gutierrez
<i>Reference Systems Coordinator:</i>	Laura Notton
<i>Production Editor:</i>	Carla Freeman
<i>Copy Editors:</i>	Bonnie Freeman, Karen Wolf
<i>Typesetter:</i>	C&M Digital (P) Ltd.
<i>Proofreaders:</i>	Theresa Kay, Scott Oney
<i>Indexer:</i>	Kathy Paparchontis
<i>Cover Designer:</i>	Candice Harman

CONTENTS

Preface	xi
About the Editors	xiii
About the Contributors	xiv
 PART I. GENERAL APPROACHES IN POLITICAL SCIENCE	
1. History of the Discipline	3
<i>Ellen Grigsby, University of New Mexico</i>	
2. Postmodernism	11
<i>Robert Swan, Western Oregon University</i>	
3. Neoinstitutionalism	22
<i>Mark C. Miller, Clark University</i>	
4. Systemism	29
<i>Hyunsun Choi, University of North Florida</i>	
5. Rationality and Rational Choice	34
<i>Richard A. Wandling, Eastern Illinois University</i>	
6. Principal-Agent Theory	43
<i>John James Quinn, Truman State University</i>	
7. Political Psychology	51
<i>David Patrick Houghton, University of Central Florida</i>	
8. Straussians	60
<i>Scot J. Zentner, California State University, San Bernardino</i>	
 PART II. COMPARATIVE POLITICS	
9. Systems Theory and Structural Functionalism	71
<i>John R. Fisher, Northwest Missouri State University</i>	
10. Political Development and Modernization	81
<i>Sarah Wilson Sokhey, The Ohio State University</i>	
11. Statism	91
<i>Anna Batta, University of North Texas</i>	
12. Dependency and Development	99
<i>Frank A. Boyd Jr., Illinois Wesleyan University</i>	

13. Civil Wars	107
<i>M. Israel Stephens, University of North Texas</i>	
14. Terrorism	115
<i>Erik S. Case, University of North Texas</i>	
15. Political and Military Coups	124
<i>Christopher D. Moore, Bethel University</i>	
16. Resource Scarcity and Political Conflict	133
<i>Walter Thomas Casey II, University of Texas at Dallas</i>	
17. Ethnic Conflict	141
<i>Tina Kempin Reuter, Christopher Newport University</i>	
18. Comparative Political Parties: Systems and Organizations	150
<i>Alexandra Cole, California State University, Northridge</i>	
19. Electoral Systems in Comparative Perspective	159
<i>William M. Downs, Georgia State University</i>	
20. Comparative Federalism, Confederalism, Unitary Systems	168
<i>William M. Downs, Georgia State University</i>	
21. Presidentialism Versus Parliamentarism	177
<i>Joseph W. Robbins, Shepherd University</i>	
22. Comparative Judicial Politics	186
<i>David L. Weiden, Indiana University–Purdue University Indianapolis</i>	
23. Civil Society	193
<i>Ryan Salzman, University of North Texas</i>	
24. Political Culture	201
<i>Marc Hooghe, University of Leuven</i>	
25. Religion and Comparative Politics	209
<i>Bryan A. S. Rasmussen, Collin College</i>	
26. Ethnic and Identity Politics	217
<i>Dwight Vick, West Texas A&M University</i>	
<i>John T. Ishiyama, University of North Texas</i>	
27. Social Movements	225
<i>Marc Hooghe, University of Leuven</i>	
28. Gender and Politics	233
<i>Karen J. Vogel, Hamline University</i>	
29. Comparative Environmental Politics and Conflict	242
<i>Amanda M. Rosen, Webster University</i>	
30. Totalitarianism and Authoritarianism	249
<i>Edward Webb, Dickinson College</i>	
31. Semiauthoritarianism	258
<i>Christian Göbel, Lund University</i>	
32. Models of Democracy	267
<i>John T. Ishiyama, Tatyana Kelman, and Anna Pechenina, University of North Texas</i>	
33. Processes of Democratization	275
<i>Eric M. Jepsen, University of South Dakota</i>	
34. Comparative Methods	283
<i>Michelle C. Flores, University of Southern California</i>	
35. Case Studies	293
<i>Catherine L. Johnson, University of North Texas</i>	

PART III. INTERNATIONAL RELATIONS

36. History of International Relations	303
<i>Anthony D. Lott, St. Olaf College</i>	
37. Realism and Neorealism	311
<i>Svetla Ben Itzhak, University of Kansas</i>	
38. Idealism and Liberalism	319
<i>Eric A. Heinze, University of Oklahoma</i>	
<i>Brandy J. Jolliff, University of Colorado at Boulder</i>	
39. Dependency and World-Systems	327
<i>Gloria C. Walker, Centenary College of Louisiana</i>	
40. Foreign Policy Analysis	336
<i>Baris Kesgin, University of Kansas</i>	
41. Feminist International Relations	344
<i>Karen J. Vogel, Hamline University</i>	
42. Leadership and Decision Making	353
<i>Tobias Van Assche, University of Antwerp</i>	
43. Balance of Power	361
<i>Brian R. King, Muskingum University</i>	
44. Deterrence Theory	368
<i>Seth M. Goldstein, The Ohio State University</i>	
<i>Brian R. King and Richard Arnold, Muskingum University</i>	
45. Rivalry, Conflict, and Interstate War	376
<i>Eric Cox, Texas Christian University</i>	
46. The Democratic Peace	384
<i>Steven R. Liebel, University of North Texas</i>	
47. Global Politics of Resources and Rentierism	393
<i>Giray Sadik, Gazi University</i>	
48. Complex Interdependence and Globalization	401
<i>Djamel Eddine Laouisset, Canada Consulting Group, Inc.</i>	
49. International Political Economy and Trade	407
<i>Stefan Fritsch, Bowling Green State University</i>	
50. Nonstate Actors in International Relations	414
<i>Michael Kuchinsky, Gardner Webb University</i>	
51. International Organizations and Regimes	422
<i>Byungwon Woo, The Ohio State University</i>	
52. International Law	431
<i>Tina Kempin Reuter, Christopher Newport University</i>	
53. International Environmental Politics	440
<i>Jerald Mast, Carthage College</i>	

PART IV. POLITICAL SCIENCE METHODOLOGY

54. Evolution of Science in Political Science 451
James K. L. Wong, London School of Economics and Political Science
55. Positivism and Its Critique 459
William Rose, Albion College
56. Constructivism 470
Mariya Y. Omelicheva, University of Kansas
57. Regression Analysis 478
Mark Zachary Taylor, Georgia Institute of Technology
58. Content Analysis 490
Marijke Breuning, University of North Texas
59. Longitudinal Analysis 497
Shane Nordyke, University of South Dakota
60. Qualitative Versus Quantitative Research 506
Alicia Jencik, University of New Orleans
61. Survey Research 514
H. Whitt Kilburn, Grand Valley State University
62. Experiments 523
Elizabeth A. Bennion, Indiana University South Bend
63. Formal Theory and Spatial Modeling 532
Donald M. Gooch, Arkansas Tech University
64. Game Theory 541
Byungwon Woo, The Ohio State University

PART V. POLITICAL THOUGHT

65. The Ancients 553
Steve G. Hartlaub, Frostburg State University
66. Asian Political Thought 560
Timothy J. Lomperis, Saint Louis University
67. Islamic Political Thought 568
Svetla Ben Itzhak, University of Kansas
68. Christian Political Thought 577
Michelle D. Deardorff, Jackson State University
69. Early Moderns and Classical Liberals 587
David Schultz, Hamline University
70. Neoclassical Liberals 596
Ellen Grigsby, University of New Mexico

71. Modern Democratic Thought	605
<i>Jeffrey D. Hilmer, University of British Columbia</i>	
72. Modern Liberalism, Conservatism, and Libertarianism	615
<i>Johnny Goldfinger, Marian University</i>	
73. Anarchism	625
<i>Leonard Williams, Manchester College</i>	
74. Nationalism	633
<i>Salih Bicakci, Isik University</i>	
75. Fascism and National Socialism	639
<i>Gerald L. McCallister, University of North Texas</i>	
76. Marxism	648
<i>Michael D'Amore, Sussex County Community College</i>	
<i>John T. Ishiyama, University of North Texas</i>	
77. Revisionism and Social Democracy	655
<i>Jutta A. Helm, Western Illinois University</i>	
78. Leninism, Communism, Stalinism, and Maoism	664
<i>Anna N. Makhova Gregg, Austin Peay State University</i>	
79. Socialism in the Developing World	673
<i>Lilly Kelan Lu, University of North Texas</i>	

PART VI. AMERICAN POLITICS

80. Founding of the American Political System	685
<i>Richard J. Hardy, Western Illinois University</i>	
81. Urban Politics	699
<i>June Sager Speakman, Roger Williams University</i>	
82. Media and Politics	708
<i>Bas W. van Doorn, The College of Wooster</i>	
83. U.S. Congress	717
<i>Tom Miles, University of North Texas</i>	
84. The Presidency	725
<i>Rebecca E. Deen, University of Texas at Arlington</i>	
85. American Judicial Politics	735
<i>Shannon Ishiyama Smithey, Westminster College</i>	
86. American Bureaucracy	743
<i>William J. Miller, Southeast Missouri State University</i>	
<i>Julia Beckett, The University of Akron</i>	
87. Interest Groups and Pluralism	752
<i>Stephen Sussman, Barry University</i>	
88. American Federalism	760
<i>Brian P. Janiske, California State University, San Bernardino</i>	
89. American Political Parties	769
<i>Jeffrey Kraus, Wagner College</i>	
90. State and Local Government	779
<i>June Sager Speakman, Roger Williams University</i>	
91. Public Policy and Administration	786
<i>Candace C. Young, Truman State University</i>	
92. Campaigns	796
<i>Chapman Rackaway, Fort Hays State University</i>	
93. Political Socialization	805
<i>Vanessa P. Bouché, The Ohio State University</i>	

94. Voting Behavior	813
<i>C. J. Williams, University of North Texas</i>	
95. American Foreign Policy	822
<i>Jeremy R. Backstrom, University of North Texas</i>	
96. Race, Ethnicity, and Politics	831
<i>Anthony DeForest Molina, University of South Dakota</i>	
97. Gender and Politics in the United States	840
<i>Angela L. Bos, College of Wooster</i>	
<i>Paula L. O'Loughlin, University of Minnesota, Morris</i>	
98. Religion and Politics in America	847
<i>Bryan A. S. Rasmussen, Collin College</i>	
99. LGBT Issues and the Queer Approach	856
<i>Robert Kulpa, Birkbeck College</i>	
Index	865

PREFACE

These two volumes of *21st Century Political Science: A Reference Handbook* were the product of many discussions that we, the editors, have had over the years concerning how to make fairly complex approaches in political science accessible to advanced undergraduate students and beginning graduate students. There is very little in the way of reference works in political science that are sufficiently accessible that undergraduate students can profitably use them to assist the pursuit of their research interests. In particular, we have sought to produce a single work that would provide students with the essentials of various approaches (both theoretical and methodological) in political science. Needless to say, our focus on essentials has meant covering fairly broad areas in the discipline, rather than specific topics. In our view, this broad focus would be most useful to undergraduate students.

In consultation with our editorial advisory board, made up of a number of eminent scholars from a variety of different subfields (who are also award-winning teachers), we selected 99 of the most important general topics in the discipline. Via these 99 entries or chapters, the SAGE 21st Century Reference Series volumes on political science highlight the most important topics, issues, questions, and debates that any student obtaining a degree in this field ought to have mastered for effectiveness in the 21st century. The purpose is to provide undergraduate majors in political science with an authoritative reference source that will serve their research needs with far more detailed information than short encyclopedia entries but not so much jargon, detail, or density as a journal article or a research handbook chapter.

To accomplish these goals, the two volumes are divided into six major parts: (I) General Approaches in Political Science, (II) Comparative Politics, (III) International Relations, (IV) Political Science Methodology, (V) Political Thought, and (VI) American Politics. In Part I, we cover the history of the discipline (e.g., the behavioral revolution, the rise of neoinstitutionalism, and the postbehavioral critique), as well as several general approaches in political science (such as rational choice, political psychology, and principal-agent theory).

Part II, on comparative politics, focuses on topics related to political development (such as modernization theory, dependency and development, statism), political violence (e.g., coups, civil wars, terrorism, ethnic conflict), political institutions (the effects of electoral laws, presidentialism, federalism, comparative judicial politics), political culture and civil society (religion and comparative politics, ethnic identity), and comparative methods (case studies, most-similar and most-different systems approaches).

Part III deals with essential approaches in international relations, including chapters on realism and neorealism, liberalism, world-systems analysis, and foreign policy analysis. There are also chapters on international conflict and war (e.g., on the balance of power, rivalry and interstate war, and the democratic peace), international political economy (e.g., complex interdependence, trade, and resource scarcity and rentierism), and global governance (international organizations and regimes and international law).

Political science methodology is covered by Part IV. We begin with chapters on the philosophy of science (including empirical approaches, positivism and its critique, and constructivism), followed by chapters that illustrate commonly used quantitative and qualitative techniques (such as regression analysis, survey research methods, experimentation, and content analysis) and then by chapters on game theory and formal modeling approaches in political science. These chapters in particular are meant to be easily understandable to students who are just beginning to engage in political science research.

Part V includes chapters on political thought, not only Western political thought but from elsewhere in the world as well. We made a conscious effort to include chapters not only on the Western classics (the “ancients,” enlightenment thinkers, neoclassical liberalism, socialism, anarchism, etc.), but also on Asian political thought, Islamic political thought, and Christian political thought. Thus students will have exposure to points of view that are not entirely rooted in the Western experience. The ability to view fundamental political issues from different points of view is, we believe, an essential skill students must have for the 21st century.

Finally, in Part VI, we cover American politics. We include chapters on the political structures and institutions of the United States (including chapters that cover research on Congress, the presidency, the bureaucracy, federalism, state and local politics, and the media) and political behavior (including public opinion and voting behavior, as well as policy making and administration). Furthermore, we have included a section on an increasingly important area in the study of American politics (which we believe will only grow in importance in the 21st century): identity politics. There are chapters that cover topics such as race, ethnicity, and politics; gender and politics; religion and politics; and LGBT issues and queer theory.

We would like to thank our families, particularly our daughters, Fasika and Bedelwa Ishiyama, and John's son, David Ishiyama, for their constant support and patience

with Mom and Dad as we finished this project (seemingly permanently tethered to our computers). We would like to thank our editorial advisory board, Larry Baum, Janet Box-Steffensmeier, Michelle Deardorff, Kerstin Hamann, and Pat James, for their wonderful suggestions regarding the topics covered by these volumes and for their constant support and encouragement as we undertook this massive project. We would also like to thank Sanford Robinson, Jim Brace-Thompson, Laura Notton, and Yvette Pollastrini at SAGE for their professionalism, their invaluable assistance, and their patience with us as we struggled through the process. All were incredibly helpful, but we would especially like to single out Sanford as he was an invaluable ally in helping to "bring the herd in." We cannot thank you all enough.

John T. Ishiyama and Marijke Breuning

ABOUT THE EDITORS

John T. Ishiyama (PhD, Michigan State University) is professor of political science at the University of North Texas. His teaching interests include comparative political development, democratization and political parties, and political methodology. His research interests include democratization and political parties in post-Communist Russian, European, Eurasian, and African (especially Ethiopian) politics; ethnic conflict and ethnic politics; and the scholarship of teaching and learning. He is the author or editor of four books and author or coauthor of more than 100 journal articles (in such journals as the *American Political Science Review*, *Perspectives on Politics*, *PS: Political Science and Politics*, *Comparative Political Studies*, *Comparative Politics*, *Political Science Quarterly*, *Political Research Quarterly*, *Social Science Quarterly*, *Europe Asia Studies*, *Party Politics*, and others) and book chapters on democratization, party politics, ethnic politics, and post-Communist Russian, European, and African politics. He has also written widely on curriculum development, educational opportunity, and educational assessment in political science. In addition, he is a research fellow at the University of Kansas's Center for Russian and East European Studies. Currently, he serves as editor-in-chief (and was founding editor) of the *Journal of Political Science Education*, the journal of the American Political Science Association (APSA) Political Science Education (formerly the Undergraduate Education) section.

He was also a member of the APSA Executive Council (2007–2009). He is currently an executive board member of the Midwest Political Science Association (2010–2013) and a member of the executive board of Pi Sigma Alpha (the national political science honorary society) from 2008 to 2012.

Marijke Breuning (PhD, The Ohio State University) is professor of political science at the University of North Texas. Her teaching interests include international relations, comparative foreign policy, international organization, and European politics. Her research interests include comparative foreign policy, foreign aid and development cooperation, and ethnic politics. She is a member of the International Studies Association (ISA), the American Political Science Association (APSA), the International Society of Political Psychology, and the European Union Studies Association. She has served on the executive committees and is past president of the APSA Foreign Policy Division, the ISA Midwest region, and the ISA Foreign Policy Analysis section. She has published articles in, among other journals, *American Political Science Review*, *International Studies Quarterly*, *Comparative Political Studies*, and *International Politics* and has coauthored a book with John Ishiyama. Together with Dr. Ishiyama, she edits the *Journal of Political Science Education*. She was part of the original ISA editorial team of *Foreign Policy Analysis*.

ABOUT THE CONTRIBUTORS

Richard Arnold is assistant professor of political science at Muskingum University, where he teaches comparative politics and international relations. He received his BA in 2003 from York University in the United Kingdom. In 2009, he graduated from The Ohio State University with a doctorate in political science, with a focus on ethnic conflict in comparative perspective. His research interests include ethnic conflict, neo-Nazi movements, nationalism, racism, anti-Semitism, qualitative methodologies, and issues in political theory.

Jeremy R. Backstrom is a doctoral student of comparative politics, international relations, and political science at the University of North Texas. His research interests include ethnic politics, civil wars, and international relations.

Anna Batta is a PhD student at the University of North Texas focusing on comparative politics, international relations, post-Communism, and the Middle East.

Julia Beckett is an associate professor of public administration at The University of Akron. She earned a JD from Washington University in St. Louis and a PhD in public administration from the University of Colorado at Denver. Her research interests are the Hoover Commission reforms, public law, public budgeting, and emergency management.

Svetla Ben-Itzhak is a PhD student in political science at the University of Kansas. She holds an MA in political science with a certificate in international service from Kansas State University. Her research interests include international relations, foreign policy, and political economy, with a special research focus on the role of foreign aid in promoting development, abating poverty, and mitigating conflict in the Middle East and Africa.

Elizabeth A. Bennion is associate professor of political science at Indiana University South Bend. She received her BA in American studies from Smith College and her MA and PhD in political science from the University of Wisconsin–Madison. Bennion uses randomized field experiments to test the effectiveness of voter registration and mobilization techniques. She has also published work

in the areas of gender politics and the scholarship of teaching and learning.

Salih Bicakci is an assistant professor in the department of International Relations of Isik University, Istanbul, Turkey. He received his BA from the Department of History at Marmara University in 1994 and his MA from the Turkic and Turkish Studies Institute of Marmara University. He teaches courses on the Middle East in international politics (with a focus on the Israeli-Palestinian conflict, Lebanon, and Iran), theory of international relations, and Turkish foreign policy. He has field experience in Israel, Palestine, and Central Asia and has published articles on security, nationalism, and international politics. Recently he has focused on security issues of the Middle East and Central Asia, including terrorism and cyberwars.

Angela L. Bos is assistant professor of political science at the College of Wooster. She received her BA from the University of Minnesota–Morris and her MA and PhD from the University of Minnesota–Twin Cities. Her research interests are gender and racial stereotypes, political nominations and elections, civic engagement and civic education, and pedagogy in political science.

Vanessa P. Bouché is a PhD candidate at The Ohio State University. She received her master's in public policy from the Lyndon B. Johnson School of Public Affairs at the University of Texas and her BA from Columbia University. Her dissertation is on the identity mechanisms of political engagement, with significant emphasis on identity-centered socialization into the affective, cognitive, and behavioral components of political attitudes. Her research and teaching interests include American politics, identity politics, political psychology, public opinion, and American foreign policy.

Frank A. Boyd Jr. is associate provost and associate professor of political science at Illinois Wesleyan University. His research interests include the political economy of development and the political economy of common pool resources.

Erik S. Case is a graduate student at the University of North Texas at Denton. He currently holds an MS in political science and is pursuing an MS in economic research. He received his bachelor's degree in political science from St. Edwards University at Austin in 2005, and his work as a counterterrorism analyst for the Texas Department of Public Safety led him to focus on terrorism-related issues during his graduate studies.

Walter Thomas Casey II is a research associate–postdoctoral fellow in electoral behavior research at the School of Economic, Political and Policy Sciences at the University of Texas at Dallas. His work focuses on the intersection of gender and political violence, the results of harmful tax competition, and modern praetorianist problems in developed democracies.

Hyunsun Choi is assistant professor in the Department of Political Science and Public Administration at the University of North Florida. He specializes in community development and urban planning related to regional governance, smart growth, housing, faith initiatives, minority communities, and public policy and, as minor fields, comparative and international development and geographic information systems. He received his PhD and a master's degree in planning from the University of Southern California, Los Angeles, and his BA and MA in public administration from Yonsei University, Seoul, South Korea. He has published peer-reviewed articles, book chapters, and professional reports and has presented a number of papers at national and international academic conferences.

Alexandra Cole is an associate professor of political science at California State University, Northridge. She received her PhD from the University of California at Irvine, and her interests include political parties and party systems. She has written about parties of the far right in Europe and most recently about changes in the German political party system.

Eric Cox received his BA from Austin College in Sherman, Texas, in 1997 and his MA (2000) and PhD (2006) from the University of Virginia in Charlottesville. He is currently an assistant professor at Texas Christian University. His research interests include the resolution of enduring rivalries, diplomacy at the United Nations, and domestic influences on the foreign policy of states.

Michael D'Amore is an Assistant Professor of Political Science and History at Sussex County Community College. He received his BA in Political Science from Moravian College in 1992, his MA in Political Science from the University of Vermont in 1998, and his PhD in Political Science from the University of Massachusetts Amherst in 2009. His dissertation was titled “The Influence of American Exceptionalist Thought on the Role of the United States in the Contemporary War on Terrorism.”

Michelle D. Deardorff is associate professor in the Department of Political Science at Jackson State

University in Jackson, Mississippi. Since earning her PhD from Miami University in 1993, her teaching and research have focused on the constitutional protections surrounding gender, race, and religion, as well as on political theory. Prior to her relocation to Jackson State in 2003, she was the Griswold Distinguished Professor of Political Science at Millikin University in Decatur, Illinois. In 2009, *Assessment in Political Science*, a primer on programmatic and classroom assessment within the discipline, which she coedited with Kerstin Hamann and John Ishiyama, was published. The two-volume set *Constitutional Law in Contemporary America*, which she coauthored with David Schultz and John R. Vile, is being published in 2010. She has been teaching political philosophy at the graduate and undergraduate levels since 2004.

Rebecca E. Deen earned her doctorate in political science from The Ohio State University in 1997. She is an associate professor of political science at the University of Texas at Arlington and chair of the Department of Political Science. Her publication and research fields include the American presidency and pedagogies in political science. She has published articles in *Congress & The Presidency*, *Presidential Studies Quarterly*, *Judicature*, *PS: Political Science & Politics*, and *Spectrum: The Journal of State Government*. The George Bush Presidential Library Foundation and the Gerald R. Ford Foundation have supported her work with travel grants. She serves as a reviewer for many scholarly journals and publishers.

William M. Downs (PhD, Emory University) is chair of the Department of Political Science at Georgia State University and codirector of the Center for Human Rights and Democracy. His research and teaching interests are in the field of comparative politics, with a specific regional focus on Europe and the member states of the European Union. He has particular research foci on parties in legislatures, federalist remedies to regional nationalism, coalition behavior, and extremist political parties. He is the author of *Coalition Government, Subnational Style: Multiparty Politics in Europe's Regional Parliaments*, which investigates the machinations of postelection alliance building in the subnational legislatures of Germany, France, and Belgium. He is editor of as well as contributor to “Regionalism in the European Union,” a special issue of *The Journal of European Integration*, and has published research articles on political parties in a long list of scholarly journals. He is coeditor of *e Extreme*, a newsletter of the European Consortium of Political Research's Standing Group on Extremism and Democracy.

John R. Fisher is an assistant professor of communications at Northwest Missouri State University in Maryville, Missouri. He received his PhD from the University of Alberta (Canada), where he studied the impact of mass communication on political decision making. His current research interests include crisis communication and media bias.

Michelle C. Flores is a PhD candidate in the Department of Political Science at the University of Southern California and a visiting instructor at the University of Northern Iowa. Her dissertation analyzes institutional change in Lebanon's state-recognized religious institutions.

Stefan Fritsch is an assistant professor of political science at Bowling Green State University. His research focuses on global political economy, technology and global affairs, multinational corporations, and European integration. Recent publications have appeared in *Global Society*, *Review of International Political Economy*, and *International Studies Perspectives*. His current research focuses on fair trade, multinational corporations, and technological evolution and the global governance of technological standards.

Christian Göbel is research fellow at the Centre for East and South-East Asian Studies at Lund University. He received his PhD from the University of Duisburg-Essen. Göbel is the author of *The Politics of Rural Reform in China: State Policy and Village Predicament in the Early 2000s* and is currently leading a research project on the consolidation of authoritarian regimes. His research interests include the change and persistence of democratic and authoritarian systems, central-local relations and governance in China's multitiered administration, and the comparative study of corruption.

Johnny Goldfinger is an assistant professor of political science at Marian University. His research interests include deliberative democracy, political engagement, and social choice theory. He teaches courses on the history of political theory and democratic decision making.

Seth M. Goldstein is a PhD candidate in political science at The Ohio State University, where he specializes in international security, political psychology, and foreign policy. He received his BA in 2003 from the University of Rochester. His research interests include signaling and perception in coercive diplomacy, the effects of domestic politics and leaders on foreign policy credibility, the psychology of threat perception in the post-September 11 era, the effect of group ideologies on risk taking in foreign policy decision making, and contemporary issues in American political life.

Donald M. Gooch is an assistant professor of political science in the Department of History and Political Science at Arkansas Tech University. He earned his PhD in political science from the University of Missouri, Columbia, in 2009. Dr. Gooch is an expert in American politics and public policy. His current research agenda involves studying trends in political polarization in the American electorate and developing empirical tests of formal voting models consistent with theoretical expectations.

Ellen Grigsby (PhD, University of North Carolina, 1987) teaches political theory at the University of New Mexico. She is the author of *Analyzing Politics* (4th ed.).

Richard J. Hardy is professor and chair of the Department of Political Science at Western Illinois University. He earned his BA from Western Illinois University, MA from the University of North Dakota, and PhD from the University of Iowa. He has served on the faculties of Northern State University, Duke University, and the University of Missouri-Columbia and has taught courses in American politics, constitutional law, civil rights, the Supreme Court, federalism, state government, policy evaluation methods, and civic leadership. He has received more than 50 significant teaching and civic leadership awards, has conducted nearly 100 workshops on constitutional law, was a member of the Electoral College, and has served as a political analyst for numerous media outlets. He has published or edited seven books, including *Government in America* and *Civic Education*, and his articles have appeared in the *American Journal of Political Science*, *Law and Policy Quarterly*, and *Publius*.

Steve G. Hartlaub completed his PhD at Michigan State University and has taught at Frostburg State University for the past 15 years. He has taught a number of honors seminars relating aspects of ancient political philosophy to modern society, including seminars on education, the regulation of sexuality, the use of rhetoric, and the understanding of tyranny. He has published a number of articles on teaching in political science and presented conference papers relating Aristotle's best regime to new urbanism, and reading Machiavelli's *Prince* from the perspective of Aristotle's teaching on tyranny.

Eric A. Heinze is assistant professor of political science and international studies at the University of Oklahoma. He holds a PhD in political science from the University of Nebraska, where he studied international relations and political theory. His current and recent research is on humanitarian intervention, liberal international theory, global governance, nonstate actors in armed conflict, and the legal and ethical implications of the war on terrorism. He is the author of *Waging Humanitarian War: The Ethics, Law and Politics of Humanitarian Intervention* (and coeditor of *Ethics, Authority and War: Non State Actors and the Just War Tradition*).

Jutta A. Helm earned her PhD at the University of North Carolina. She is currently professor emerita of political science at Western Illinois University. Her teaching interests include European politics, German politics, and the comparative study of genocide. Most of her publications focus on political participation, broadly defined: new social movements, workplace democracy, and the role of trade unions in hard times.

Jeffrey D. Hilmer is a postdoctoral teaching fellow at the University of British Columbia, Vancouver. He received his PhD in political science, with an emphasis on democratic theory, from the University at Albany, State University of New York, in 2007. His current research interests include differentiating participatory and deliberative theories of

democracy, exploring participatory democratic institutions in Brazil, theorizing agonistic democratic institutions, and contrasting old and new theories of pluralism.

Marc Hooghe is a professor of political science at the University of Leuven (Belgium). He obtained his PhD in political science in 1997 from the University of Brussels (Belgium), with a dissertation on the political impact of new social movements. His work focuses mainly on political participation, social movements, and social capital, and it has appeared in various scientific journals, such as the *European Journal for Political Research*, *Political Behavior*, and *Comparative Political Studies*. At Leuven University, he is head of the Centre for Political Science Research.

David Patrick Houghton is an associate professor of political science at the University of Central Florida. His published books include *U.S. Foreign Policy and the Iran Hostage Crisis* and *Political Psychology: Situations, Individuals, and Cases*. His latest book, *The Decision Point*, is due to be published in 2011. His area of expertise is political psychology and decision making in international relations, and he has published refereed articles in a variety of top journals. He received his PhD in political science in 1996 from the University of Pittsburgh and has also taught there and at the University of Essex. From 2001 to 2002, he was a visiting scholar at the Mershon Center for International Security Studies at The Ohio State University.

Brian P. Janiskee is professor of political science and department chair at California State University, San Bernardino. He received his PhD from Michigan State University. Janiskee is coauthor of *Democracy in California* (2008), coeditor of *The California Republic* (2003), and author of numerous articles in such areas as crime policy, local government, and presidential elections. His current research interests include local government in the American colonial period, research methodology, and national security policy.

Alicia Jencik is a PhD candidate in the department of political science at the University of New Orleans. She specializes in political behavior, gender politics, and political institutions. Her dissertation examines how women and men are affected differently by natural disasters and uses the city of New Orleans before and after Hurricane Katrina as a case study. She obtained her BA, MA, and a certificate in women's studies at Florida Atlantic University. She is the author of *A Woman in the White House? The Effects of Gender on Female Politicians*, as well as "Survey Research in Post Katrina New Orleans: 'I Can't Imagine It,'" with Susan Howell. She worked with the University of New Orleans Survey Research Center from 2005 through 2008, conducting research on quality of life and recovery issues in post-Hurricane Katrina New Orleans. She has focused on survey research methodology and has taught research methods courses to undergraduate students.

Eric M. Jepsen is an assistant professor at the University of South Dakota. His teaching and research interests include political economy, democratization, globalization, Latin American politics, and South Asian politics. His research on subnational economic development and democracy in India and Mexico has been funded by the National Science Foundation and the Social Science Research Council and has appeared in a number of books and journals.

Catherine L. Johnson is a PhD candidate and teaching fellow in the Department of Political Science at the University of North Texas. She earned an MPA from the University of North Texas in 2003. Her dissertation topic involves the influences of military service on political attitudes, voting, and participation.

Brandy J. Joliff is a PhD student in political science at the University of Colorado, where she is studying comparative politics and international relations. She holds an MA in international studies from the University of Oklahoma. Her current work deals with comparative European foreign policy, European integration, and democracy promotion to third world countries in European Union foreign policy.

Tatyana Kelman is a graduate student in the PhD program in political science at the University of North Texas.

Baris Kesgin is a PhD candidate in political science at the University of Kansas. His work (with Juliet Kaarbo) has appeared in *International Studies Perspectives*. Presently, his research concentrates on tracing leadership effects in Israel's and Turkey's relations with the United States.

H. Whitt Kilburn is assistant professor of political science at Grand Valley State University. He received his PhD from the University of North Carolina at Chapel Hill in 2005. His research interests center on the study of mass public opinion, elections, and voting behavior.

Brian R. King is assistant professor of political science at Muskingum University. He received his BA in 1992 from Ohio Northern University and his PhD in 2005 from the University of Cincinnati, where he studied both American politics and international relations. His research interests include presidential management and policy, organizational learning in government, American political institutions, international security issues, international crisis decision making and management, and issues related to American foreign policy and national security.

Jeffrey Kraus is associate provost and professor of government and politics at Wagner College, where he has been a member of the faculty since 1988. Prior to coming to Wagner, he taught at Kingsborough Community College (1980–1987) and Baruch College (1987–1988). His research has focused on political parties, elections, and campaign finance. He has also been a consultant to a number of political campaigns in New York and since 2001 has chaired the New York City Voter Assistance Commission.

Michael Kuchinsky is an assistant professor of political science at Gardner-Webb University (North Carolina), where he teaches international relations, comparative politics, political economy and development studies, and other courses in global politics. Before college teaching, he served as a program associate and executive in several public policy nongovernmental organizations, as well as being an ordained clergyman.

Robert Kulpa is a PhD candidate at Birkbeck College, London, United Kingdom. His research focus is on the formation of gay identities and their interwoven connections with a national identity in post-Communist Poland. He has an interest in a comparative study of Western and Central European lesbian, gay, bisexual, and transgender movements, queer theories, nationalism, and links between power and knowledge. He teaches politics at University College London and has published in a number of scholarly journals.

Djamel Eddine Laouiset earned his PhD in 1983 from the University of Miami (Florida). He is CEO of Canada Consulting Group in Ottawa and was director of studies and research at the National Institute for Global Strategic Studies in Algiers, Algeria, from 1987 to 1990. He is also a former member of the National Research Council in Algiers and advised the minister of economy and commerce of the United Arab Emirates from 2001 to 2004. He is a fellow of the Salzburg Seminar (Austria) and a Gerson Lehrman Group council member. He has published several books, papers, and articles in his fields of expertise and has spoken at various international academic and professional forums. He is a founding and active member of a number of international academic and professional societies.

Steven R. Liebel is a PhD candidate in international relations and political science at the University of North Texas. His research interests include conflict management and processes, and human rights.

Timothy J. Lomperis is professor of political science at Saint Louis University in Missouri. He grew up in India as the son of missionaries. He received his PhD from Duke University and has taught at Louisiana State University, Duke University, the U.S. Military Academy at West Point, and Saint Louis University. He has published four major books, *The Hindu Influence on Greek Philosophy* (1984) and three books on the Vietnam War: *The War Everyone Lost—and Won* (1984, 1986, and rev. ed., 1993), *Reading the Wind* (1987), and *From People's War to People's Rule* (1996). He has published numerous articles on Vietnam and on the War on Terror. His current research interests are Asian political thought and the foreign policy of comparative military interventions.

Anthony D. Lott is an associate professor of political science at St. Olaf College in Northfield, Minnesota. He earned his PhD in international studies in 2002 from the Graduate School of International Studies at the University of Denver. His research interests include developments in

international relations theory, the history of international law and cooperation, and challenges to international security. He is the author of *Creating Insecurity: Realism, Constructivism and U.S. Security Policy* and numerous articles on international relations theory and international law.

Lilly Kelan Lu is a PhD student at the University of North Texas. She received her BA in law in 2005 from Qingdao University, and in 2007 she completed a master's degree in international affairs and administration at Missouri State University. Her research interests include comparative politics and economy, as well as the application of quantitative methodology.

Anna N. Makhova-Gregg is an assistant professor of political science at Austin Peay State University, Clarksville, Tennessee. She received her PhD from Southern Illinois University, Carbondale (2008). Her current research interests include Russian and post-Soviet politics, Chinese politics, bureaucracy in post-Communist states, and the political economy of transition to capitalism.

Jerald Mast has a bachelor's degree in political science from the University of Wisconsin–Madison and a PhD in political science from Northern Arizona University. He has conducted research on economic valuation of environmental aesthetics, environmental security, and natural resource policy on the Great Lakes. He is the chair of the political science department at Carthage College.

Gerald L. McCallister is a PhD student of political science at the University of North Texas. He received his MA in 2006 from the University of Texas at Arlington. His research interests include China–Taiwan relations, international conflict, international trade, the former Soviet Union, Mexican politics, diplomacy, terrorist groups, and U.S. foreign policy.

Tom Miles is a PhD student at the University of North Texas and holds a master's degree in political science from Texas State University. He specializes in American politics and has done extensive research on the interactions between Congress and the president. His research, with Matthew Eshbaugh-Soha (University of North Texas), on the congressional agenda setting of President George W. Bush, has been published in *American Review of Politics*.

Mark C. Miller is a professor of American government and director of the Law and Society Program at Clark University in Worcester, Massachusetts. He received his PhD from The Ohio State University and his JD from The George Washington University. He is the author of *The High Priests of American Politics: The Role of Lawyers in American Political Institutions* (1995), coeditor of *Making Policy, Making Law: An Interbranch Perspective* (2004), editor of *Exploring Judicial Politics* (2008), and author of *The View of the Courts From the Hill: Interactions Between Congress and the Federal Judiciary* (2009). His current research interests include the interactions between legislatures and courts in the United States and abroad, as well as the role of lawyers in American political institutions.

William J. Miller is an assistant professor of political science at Southeast Missouri State University. He earned his PhD in public administration and urban studies at The University of Akron. His research focuses on comparative public opinion regarding various public policies, the pedagogy of political science, and public perceptions of bureaucracy.

Anthony DeForest Molina, PhD, is director of the Master of Science in Administration program and assistant professor in the Department of Political Science at the University of South Dakota. He earned his doctorate in urban studies and public affairs at Cleveland State University, where he taught courses in urban theory, American government, and urban politics and the African American experience. His current teaching and research interests include public administration ethics, organization theory, and the role of values in administrative decision making.

Christopher D. Moore is an assistant professor of political science at Bethel University, St. Paul, Minnesota. He received his PhD from The Ohio State University in 2008. His research interests include foreign policy decision making, political psychology, and nonstate political violence, including terrorism and insurgency.

Shane Nordyke is assistant professor of political science at the University of South Dakota. She received her PhD in public policy from Indiana University in 2008. Her research interests include a variety of public policy problems but focus on those related to homeland security, terrorism, aviation security, aviation safety, and intergovernmental relations. Most recently she has completed a longitudinal analysis of spending by the Department of Homeland Security at the state level from 2002 to 2007. Currently she enjoys teaching courses in public policy, policy analysis, and research methods to students in undergraduate, MPA, and PhD programs.

Paula L. O'Loughlin is professor of political science at the University of Minnesota–Morris. She received her BA from Smith College and completed her PhD at the University of Minnesota–Twin Cities. Her research interests include gender and politics, civic education, political behavior, media, and political psychology.

Mariya Y. Omelicheva is an assistant professor of political science at the University of Kansas. She holds a PhD (2007) in political science from Purdue University and a JD in international law (2000) from Moscow's National Law Academy. Her research and teaching interests include international and Eurasian security, counterterrorism and human rights, and Russia's foreign policy. She has also published on the issues of students' assessment and application of various research techniques to the analysis of learning outcomes.

Anna Pechenina is a graduate student in the PhD program in political science at the University of North Texas.

John James Quinn is a professor of political science at Truman State University, where he teaches classes on comparative politics, African politics, international relations, international political economy, and methodology. His recent publications have centered on divergent aspects of the political economy of African development, African foreign policy, African political parties, corruption and its tie to majority state ownership, and the links between the Rwandan genocide and the later disorder in Zaire (Democratic Republic of the Congo). He is author of *The Road Oft Traveled: Development Policies and Majority State Ownership of Industry in Africa* (2002).

Chapman Rackaway is associate professor of political science, Fort Hays State University. He received his PhD from the University of Missouri in 2002. His specialty is in the area of campaign processes and political organizations, specifically the intersection of political parties and their affiliated campaigns at the state and local levels. His work has been featured in *Politicking Online* (2009, Rutgers University Press) and in the *Social Science Computer Review*.

Bryan A. S. Rasmussen is a professor of political science at Collin College in Frisco, Texas. He received his MA from the University of Nevada, Las Vegas. He has continued his graduate work and his study of religion and politics at the University of North Texas in Denton. He has a passion for teaching undergraduates the importance of their government.

Tina Kempin Reuter, PhD, is assistant professor of international and comparative politics at Christopher Newport University, Newport News, Virginia, and director of the program in international conflict management. She received her MA (2002) in contemporary history, economics, and international law and her PhD (2006) in international law and international relations, both with distinction, from the University of Zurich in Switzerland. She was a visiting researcher at the Solomon Asch Center for Study of Ethnopolitical Conflict at the University of Pennsylvania (2005–2006), a research fellow at the Institute of Public International Law at the University of Zurich Law School (2003–2005), and a research assistant at the Center for Security Studies and Conflict Research at the Swiss Federal Institute of Technology (1999–2002). She is the author of *Ethnic Conflict and International Law: Group Claims and Conflict Resolution Within the International Legal System* (forthcoming) and *Ready for Peace? The Implementation of the Good Friday Agreement in Northern Ireland, 1998–2002* (2003). She has also published various articles on ethnic conflict, human rights, and minority issues.

Joseph W. Robbins is currently an assistant professor in the Department of Political Science at Shepherd University. He received his PhD from the Department of Political Science at Texas Tech University in 2008. His

work has focused on fiscal policy making, political institutions, and the development of political party systems in former Communist countries. His work has appeared in a number of journals, including *Comparative Politics*, *Party Politics*, and *Europe Asia Studies*.

William Rose is associate professor and chair of the Department of Political Science at Albion College in Michigan. He holds a JD from the University of Toledo and a PhD in political science from the University of Massachusetts–Amherst. His scholarly and teaching interests lie in the areas of contemporary legal and political theory, the history of American legal and political thought, and jurisprudence.

Amanda M. Rosen is an assistant professor of politics and international relations at Webster University. She earned a bachelor's degree from Duke University in political and economic studies of Europe and a master's degree and PhD (2009) in political science from The Ohio State University, where her dissertation was titled *Emission Impossible? The Impact of the International Climate Regime on Sub National Climate Change Policymaking*. In addition to climate change policy making, she is interested in American foreign policy, international security, and American political behavior. She has received several grants for her research, including a grant from the Merzhon Center for International Security Studies, and has presented papers at the Midwest Political Science Association and the Joint Sessions of the European Consortium for Political Research.

Giray Sadik earned his PhD (2008) in political science from the University of Georgia, specializing in international relations and comparative politics. His research interests encompass the interaction of domestic politics and foreign policy, focusing on areas such as American foreign policy and comparative Middle Eastern political systems. He was recently a visiting assistant professor of political science at Truman State University, teaching courses on American foreign policy and international security, and is now in the Department of International Relations at Gazi University, in Ankara, Turkey.

Ryan Salzman is a PhD candidate studying political science at the University of North Texas, Denton. His research interests include the comparative politics of political culture in developing countries. He is currently studying the effects of media consumption on democratic norms and behavior in the Latin American region.

David Schultz is professor in the School of Business at Hamline University in St. Paul, Minnesota. He is the author or editor of more than 25 books and 70 articles on various aspects of political theory, law, and American politics, including *Inventors of Ideas: An Introduction to Western Political Philosophy*.

Shannon Ishiyama Smithey is associate professor in the Department of Political Science and Sociology at Westminster College in New Wilmington, Pennsylvania. She received her PhD in political science from The Ohio State University in 1994. Her research focuses primarily on courts in countries making transitions to democracy, with particular emphasis on judicial interpretation of electoral rights, civil liberties, and court powers.

Sarah Wilson Sokhey is a PhD candidate in the Department of Political Science at The Ohio State University. In the fall of 2010, she will be an assistant professor at the University of Colorado, Boulder. Her research interests focus on comparative political economy, economic development, interest group politics, business–state relations, and the post-Communist countries of Eastern Europe, Russia, and Central Asia.

June Sager Speakman, PhD, is professor of political science at Roger Williams University in Bristol, Rhode Island, where she teaches American politics and public administration. She has a PhD in political science from City University of New York and an MA in economics from the New School for Social Research. She has done research on politics and public policy at the municipal level and is currently serving her third term of office on her local town council.

M. Israel Stephens has a MPSA from The Bush School at Texas A&M University and currently is a PhD (ABD) student at the University of North Texas. His research interests include international politics at the systemic level and conflict, broadly defined. He should finish his dissertation in 2011 and will look to continue in the ranks of academia.

Stephen Sussman is an assistant professor of public administration and academic coordinator at Barry University's School of Adult and Continuing Education, where he teaches in the MPA and BPA programs. He earned his MPA and his PhD in political science from Georgia State University. He serves on the Graduate Council and the Faculty Senate. In 2007, he was appointed to a Department of Education Federal Rulemaking Committee as a primary negotiator for regulations relating to the General Provisions for the Title IV programs of the Higher Education Act of 1965. In 2010, he coauthored a poster presentation for the International Preparedness and Response to Emergencies and Disasters conference in Tel Aviv, Israel, titled *Disaster Planning and Analysis: An Evaluation of Emergency and Disaster Preparedness Education*.

Robert Swan is an assistant professor of criminal justice at the Western Oregon University. He received his BS in political science, his MS in criminology and criminal justice, and his PhD in public administration and policy from the Hatfield School of Government at Portland State University, Portland, Oregon. His research interests include an exploration of the presence (or absence) of the new penology in the American criminal justice system; the study and use of interstate inmate

transfers by American corrections officials; decision-making influences on correctional management, especially among prison wardens and superintendents; and the impact of vicarious trauma on prosecutorial decision making among prosecutors tasked with prosecuting sex offenders.

Mark Zachary Taylor is an assistant professor at the Georgia Institute of Technology in the Sam Nunn School of International Affairs. Formerly a solid-state physicist, he earned his PhD in political science at the Massachusetts Institute of Technology. In his research, he uses regression analyses of patents, scholarly publications, and high technology production data to test hypotheses about the political economy of science, technology, and economic competitiveness. His research can be found in the journals *Foreign Affairs*, *Harvard International Review*, *International Organization*, *Journal of Political Science Education*, and *Review of Policy Research*.

Tobias Van Assche is a postdoctoral fellow at the University of Antwerp. He received his PhD in political science from Syracuse University. He has published a number of articles on political leadership and foreign policy decision making. His main focuses are leadership in international organizations and the study of groupthink and sequential decision making. He is also interested in the agenda-setting capabilities of political leaders.

Bas W. van Doorn is assistant professor at the College of Wooster. He received an MA in American studies from the University of Amsterdam in 2000 and a PhD in political science from the University of Minnesota in 2008. His research focuses on public opinion, political psychology, media and politics, and the presidency.

Dwight Vick is the graduate director of political science programs at West Texas A&M University, Texas A&M University system. His research focuses on issues relating to public administration and the policies surrounding addiction as well as social equity issues. Dr. Vick is the author of two books focusing on substance abuse, drug courts, and public administration. A member of the National Council for the American Society for Public Administration, he has written articles and participated in research that has been published in *Public Administration Review*, *PA Times*, *JAMA: The Journal of the American Medical Association*, and *Addiction*. He is the faculty advisor for the academic student-managed journal, *PB&J: Politics, Bureaucracy and Justice*.

Karen J. Vogel is professor of political science at Hamline University in St. Paul, Minnesota. She received her PhD from the Department of Political Science at the University of Oregon. She teaches courses in international organizations, gender politics, and European politics. Her publications include articles on feminist pedagogy and experiential education, as well as numerous conference papers on gender and European politics.

Gloria C. Walker is assistant professor of political science at Centenary College of Louisiana. She received her PhD from the University of California at Davis. Her research interests include political development, human rights, democratic consolidation and the quality of democracy, regional integration, and the role of intergovernmental and nongovernmental organizations in promoting liberalization.

Richard A. Wandling (PhD, Miami University) is a professor in the Political Science Department of Eastern Illinois University, where his primary teaching areas are public organization theory, government budgeting, and state and local politics and policy. His current research agenda focuses on environmental politics and policy making, which provides a rich set of opportunities for applying and testing alternative conceptions of political rationality and governmental decision making.

Edward Webb is assistant professor of political science and international studies at Dickinson College. After graduating from the University of Cambridge in 1992, he worked for several years in the British Diplomatic Service before leaving to pursue his PhD in political science at the University of Pennsylvania (2007). His research interests include religion and politics, cultural authoritarianism, and empires, with a focus on Turkey, the Middle East, and North Africa.

David L. Weiden is assistant professor of political science at Indiana University–Purdue University Indianapolis. He received his PhD from the University of Texas at Austin and his JD from the University of Denver College of Law. His research interests include comparative judicial decision making, the role of law clerks at the high courts of the United States and Canada, and trial courts. He is the coauthor of *Sorcerers' Apprentices: 100 Years of Law Clerks at the United States Supreme Court* (2006).

C. J. Williams is a doctoral student in political science at the University of North Texas. He received his bachelor's degree in political science from the University of Mary Washington and his master's degree in political science from the University of North Texas. His area of study is comparative politics. His main focus is political culture and its effect on individual-level attitudes and behaviors.

Leonard Williams is professor of political science at Manchester College in North Manchester, Indiana. He received his BA and MA from Southern Illinois University at Carbondale in 1974 and 1976, respectively, and his PhD from The Ohio State University in 1981. His primary research interests concern political ideologies. He has written a number of journal and popular articles on this and other topics. He also authored *American Liberalism and Ideological Change* and coedited (with Joseph Losco) *Political Theory: Classic and Contemporary Readings*.

James K. L. Wong is a PhD candidate and graduate teaching assistant at the Department of Government, London

School of Economics and Political Science. He obtained his BS (with first-class honors) in government and public administration from the Chinese University of Hong Kong in 2005 and his MS in political theory from the London School of Economics in 2006. He is specializing in judgment aggregation, democracy, and environmental political theory for his PhD thesis, and his research interests also extend to contemporary moral and political philosophy, as well as the nature and methods of political theory and political and social science.

Byungwon Woo is a PhD candidate in political science at The Ohio State University. He specializes in international organizations and international political economy and has strong game theoretic and methodological skills.

Candace C. Young has been a professor of political science at Truman State University since 1980. She holds a bachelor's

degree from Columbia College and earned her doctorate in 1982 from the University of Missouri. Her current research focuses on higher education policy, assessment of student learning, program evaluation, and legislators' use of information. She teaches courses in public policy, the presidency, and public administration and administers an internship program in Missouri state government. She has served on numerous statewide panels related to issues of higher education, including the Governor's Commission on the Future of Higher Education.

Scot J. Zentner is professor of political science at California State University, San Bernardino. He received his PhD from Michigan State University. He is the author of numerous reviews, journal articles, and book chapters. His research interests include political philosophy and American politics.

PART I

GENERAL APPROACHES IN POLITICAL SCIENCE

1

HISTORY OF THE DISCIPLINE

ELLEN GRIGSBY

University of New Mexico

Within the discipline of political science in the United States, *traditionalism*, *behavioralism*, and *postbehavioralism* are three distinct political science research approaches. That is, each offers a perspective on how best to carry out investigation, analysis, and explanation relating to politics and political life (Dryzek & Leonard, 1988). These three approaches represent different points of emphasis regarding the ways in which research about politics should proceed. For example, it will be seen that traditionalism—in comparison with behavioralism—tends to emphasize the usefulness of analyzing governmental institutions when studying political phenomena, whereas behavioralism tends to assert the importance of research into the intricacies of the behavior of individual political actors (e.g., citizens, lobbyists, candidates, elected officials). However, all three research perspectives share the belief that political science research should produce explanations that improve and deepen our understanding of complex political processes.

As one begins to analyze the meaning and complexity of traditionalism, behavioralism, and postbehavioralism, it is important to keep in mind three points. First, traditionalism, behavioralism, and postbehavioralism are broad categories, and within each category one finds a variety of political scientists who are not necessarily in agreement on all matters relating to the study of politics. For example, during the years in which traditionalism was the prevailing research approach within political science, Woodrow

Wilson (1911) delivered an address to the American Political Science Association (APSA) that called into dispute various claims made by previous APSA president James Bryce. In 1908, Bryce had stated that political science, that is, a scientific understanding of politics, was possible insofar as human actions tended to be similar, or repeatable, over time; thus, Bryce (1909) reasoned, one could generalize about patterns of human activity and draw conclusions about political life. Wilson (1911), however, while not altogether denying the existence of some degree of patterned activity over time, stressed the uniqueness characterizing human beings and human actions. Despite these differences, both Bryce and Wilson were representative of traditionalist political science.

Second, traditionalism, behavioralism, and postbehavioralism are often linked with certain decades in the development of political science in the United States. Traditionalism is usually associated with the political science practiced during the 19th and early 20th centuries. Behavioralism is generally associated with the post-World War II period, although its origins are sometimes traced back to the 1920s. Postbehavioralism's appearance in the discipline had been noted and commented on by the end of the 1960s (Dahl, 1992; Dryzek, 2006; Ricci, 1984).

It is important to realize, however, that these historical markers are best used as general designations, because the development of these three research approaches was too multifaceted and complex to fit neatly into rigid time

4 • GENERAL APPROACHES IN POLITICAL SCIENCE

categories. The emergence of a new approach did not necessarily completely or entirely displace an older one; for example, while traditionalism was challenged by behavioralism in the 1950s and 1960s, a number of political scientists continued to hold to traditionalism. Indeed, many contemporary introductory textbooks in U.S. politics continue to reflect the perspective of traditionalist political science. Moreover, not all subfields of political science were affected equally or simultaneously by the emergence of a new approach. For instance, the subfield of U.S. politics incorporated the behavioralist approach earlier than did the subfields of international relations and comparative politics (Sigelman, 2006).

Third, two of the three research approaches have tended to define themselves in opposition to their predecessors and, in so doing, have helped shape the manner in which those prior approaches have been remembered. Specifically, behavioralism defined itself in opposition to what it understood as constituting traditionalism, and post-behavioralism carved out its own identity, in part, as a critique of what it saw as the defining elements of behavioralism. As a result, one sees that the emergence of the newer approaches was coupled with a rejection of perceived deficiencies in the earlier approaches. In identifying what they saw as inadequacies in the older approaches, the newer approaches tended to highlight differences between the new and the old and, in some cases, tended to understate any similarities. For example, behavioralism emphasized its adherence to scientific method and, in so doing, sometimes gave the impression that that which it was attempting to replace—traditionalism—had not regarded itself as scientific. As becomes clear when one analyzes the actual writings of traditionalists, however, traditionalists generally saw themselves as political *scientists* and often made much of the fact that, as political scientists, they were not to be confused with historians (Farr, 1990; Gunnell, 2006). As early as 1910, an APSA president was calling on the discipline to employ statistical analyses to identify political patterns and test conclusions relating thereto (Lowell, 1910). Similarly, postbehavioralists, it will be seen in the discussion below, emphasized the importance of producing research that was relevant in addressing contemporary questions, but, in stressing their own newness relative to behavioralists, postbehavioralists often tended to understate the extent to which early-20th-century political scientists had also sought to use political science research to address urgent, relevant problems in U.S. life (Gunnell, 2006).

Traditionalism

Definition and Overview

Traditionalism is an approach defined by its focus on the study of political institutions, law, or a combination of these. In addition, traditionalism locates its scientific reliability in its grounding in careful historical or legal

investigations that are designed to produce thorough descriptions of the subject in question (Easton, 1971; Fried, 2006; Isaak, 1985; Macridis, 1992). That is, traditionalism is an approach in political science that seeks to study political phenomena by investigating law, history, and/or institutions such as the government as a whole or narrower institutions such as legislative, executive, or judicial bodies. A traditionalist seeking to understand how the U.S. Congress works would, thus, investigate such questions as what the law (e.g., the U.S. Constitution) provides for in terms of congressional powers and limits, how Congress as an institution has evolved historically, and how Congress as an institution fits into the larger institutional network of the U.S. government in its entirety. A traditionalist seeking to understand courts could follow a similar strategy of pursuing historical questions (e.g., how courts have evolved), legal questions (e.g., what laws govern courts and how courts have participated historically in shaping laws), or institutional questions (e.g., how courts are organized and administered as institutions). A traditionalist in the field of international relations might study international law or national laws and treaties relating to interstate interactions (i.e., foreign policy).

Traditionalist political science has not been an approach that has demanded narrow or exclusive disciplinary specialization. On the contrary, early traditionalist political scientists needed to be comfortable with such fields as history or law in order to pursue their work. Francis Lieber, who, in 1857, became the first person to hold an official political science professorship in the United States, was, in actuality, a professor of both history and political science at New York's Columbia College (Farr, 1990). Traditionalism's breadth is also revealed in APSA president Albert Shaw's (1907) comments that it was possible to find numerous political scientists participating in the American Historical Association as well as in "Economic and Sociological groups" (p. 178).

Traditionalist political scientists tended to be explicit in drawing connections between political science research and service to the public interest, in whatever manner the latter might be defined by the political scientist in question. Shaw's 1907 APSA presidential address is an illustration of traditionalism's linkage of empirical-scientific and normative-ethical objectives. "I believe that there will be a very general agreement," Shaw asserted, "that this Association can render an extremely useful service to the country, without departing in the smallest degree from its scientific methods" (p. 181). Shaw went on to suggest that APSA might undertake investigative projects on problems or concerns relative to "the public benefit" (p. 181). In fact, a perusal of the early records published in *Proceedings of the American Political Science Association* and in the *Annals of the American Academy of Political and Social Science* reveals traditionalists' interests in addressing child labor, political party reform, and other public welfare questions (Addams, 1906; Richberg, 1913).

Case Studies of Traditionalism: Frank Goodnow and Woodrow Wilson

For a fuller, more detailed understanding of traditionalism, one can look in greater depth at two examples of traditionalist political science. The first is Frank Goodnow's 1904 address to the first meeting of APSA. Goodnow's address included (a) a definition of what he called political science's "scope" but not a technical definition of political science itself, (b) an examination of what political science was to have as its research focus, and (c) a closing statement about political science's relevance. An examination of these three components of his address illustrates traditionalism's salient elements of institutionalism (in the emphasis on studying the institution of the state), legalism (in the emphasis on studying law and jurisprudence), a historical perspective, and attention to the public benefits of scientific inquiry.

First, in his address, Goodnow (1904) announced that he preferred to define political science's scope (i.e., that which political science was to study) rather than attempt a definition of political science itself. Setting out to construct a technically detailed definition of the discipline *per se*, Goodnow contended, was not as productive an enterprise as determining what the discipline should have as its focus of research. He pointed to what he termed the "dangerous" possibility of defining the discipline in too limited or too expansive a manner (p. 35). He proceeded to characterize political science's scope as the investigation of states. Political scientists were neither the first nor the exclusive researchers of states, Goodnow explained, but were, rather, unique in targeting the state as a primary subject for analysis. For example, historians might study historical states and might indirectly study contemporary states, Goodnow reasoned, and economists might investigate monetary matters relating to states. However, only political scientists would have as their "main interests" the direct, detailed, "scientific" analysis of states in all their complexity. Goodnow's comments suggest that the previously noted absence of disciplinary narrowness or specialization in traditionalist political science did not have to translate into the absence of disciplinary identity. Goodnow was, in this address, identifying himself as a political scientist as opposed to a historian, even while his approach to political science would employ historical perspectives. Moreover, in identifying the institution of the state (as opposed to the behavior of individuals, for example) as the central and defining subject matter of political science, Goodnow was conveying what is generally termed the *traditionalist orientation toward institutionalism*.

Second, Goodnow (1904) framed the study of states—and thus political science as a discipline—broadly. Political science's range of investigation was to include, he argued, the study of how the "State's will" was communicated, what comprised the "State's will," and how the "State's will" was carried out. In explaining what he meant by the communication of the "State's will," Goodnow made

reference to such matters as the values conveyed through a country's political ideas or political theory, constitution, and political party platforms. Political values influenced state policies or will. The second element—the "content of the State will"—Goodnow identified as law (p. 40). Law revealed a state's meaning. Indeed, one sees how closely Goodnow's traditionalist political science was attached to the study of law when one encounters his remark that "it is very doubtful" that anyone could become a political scientist—that is, that anyone could understand states "as an object of scientific study"—without a thorough understanding of law (pp. 42–43). To understand how states carried out their "wills," Goodnow continued, one needed to study administrative law, a subject that, in the absence of political science, had been frightfully neglected, he believed. He pointed to the benefits of studying the history of English poor laws as a guide for improving public administration generally.

Finally, Goodnow (1904) closed his address by expressing hope that political science could contribute to the public good. He identified teachers and political practitioners as two groups that could benefit directly from the knowledge produced by the discipline. Moreover, in disseminating a more descriptively accurate and comprehensive understanding of states, teachers and practitioners, in their respective professional roles, could contribute to an enhanced public well-being.

An examination of Woodrow Wilson's (1911) address to the seventh annual APSA meeting offers a second opportunity for scrutinizing more carefully traditionalism's breadth, a breadth critiqued as "unscientific" by later advocates of behavioralism. Although better known as the 28th president of the United States, Wilson also served as president of APSA and, in this latter capacity, argued against a narrow, specialized conception of political science. In fact, at one point in his address, he went so far as to assert that he disliked the name *political science*, which, he claimed, implied that human interactions should be studied objectively and narrowly. He argued for the designation *politics* rather than *political science* as a more suitable name for the study of the state and "statesmanship" (pp. 10–11). Although Wilson supported a scientific approach, if by science one meant accuracy and thoroughness in one's study of political life, he argued that such study should include an examination of literature, art, and poetry and should seek to inspire "vision" and "sympathy" (pp. 2, 10, 11). His understanding of political science, one finds, could hardly be broader, in that he concluded that "nothing" that has an impact on "human life" should be termed "foreign" to the discipline (p. 2). Wilson argued that the astute student of politics should demonstrate "a Shakespearian range" (p. 10). Although Wilson's immediate influence on U.S. political science was limited (Ubertaccio & Cook, 2006), his explicit embrace of an expansive politics is illustrative of traditionalism's lack of disciplinary specialization. In addition, a comparison of his approach with that of Goodnow is helpful in reminding students of traditionalism of the approach's internal diversity.

Behavioralism

Definition and Overview

Behavioralism emerged as a criticism of traditionalism's failure, in the view of behavioralists, to offer an approach to the scientific investigation of political questions that was sufficiently rigorous to produce predictive results based on quantitatively tested data. Specifically, behavioralism's defining elements include a focus on political actors and their behavior (or attitudes and opinions), value-free science, and the study of operationalizable questions through hypothesis formulation and empirical, quantitative research (Ricci, 1984). The focus on studying political actors represented a shift away from traditionalism's concentration on the historical and legalistic study of institutions.

In turning attention to the study of political actors, many behavioralists employed survey research to compare the attitudes of voters versus nonvoters, elites versus non-elites, partisan identifiers versus independents, or other subunits of populations. Students of congressional politics could enlist behavioral approaches to shift research away from the analysis of the institutional history of legislatures to an empirical investigation of the actual behaviors of congressional officeholders, staff, or congressional committee members. Behavioralists were interested, for example, in whether members of Congress spent greater time and devoted greater resources to the actual drafting of legislation or to responding to constituency demands, campaigning for the next election, or interacting with lobbyists. Empirical observation of such behaviors devoid of normative judgments (about how voters, nonvoters, elites, masses, partisans, independents, or congressional members "should" be behaving) would, in the words of David Easton (1971), correct the traditionalist "neglect of the most obvious element, the human being" (p. 203) in the conduct of research. Moreover, not only would a "value-free" science guard against the corruption of biases associated with normative preferences, but strict adherence to the study of questions translatable into operational variables and testable hypotheses would provide a more reliable knowledge than that producible by means of traditionalism.

In a 1967 essay titled "The Current Meaning of Behavioralism," Easton (1992) summed up behavioralism as having eight interrelated "intellectual foundation stones" (p. 47):

- "regularities": A rigorous study of political behavior would allow political scientists to make predictions, just as natural scientists could make predictive statements.
- "verification": Predictions were to be testable in order to be falsified or verified.
- "techniques": Political science should become increasingly sophisticated in its use of scientific data collection and testing methods.

- "quantification": Political science should use precise, quantifiable measurements; questions for research had to be definable in testable, operationally narrow and precise terms.
- "values": Empirical, scientific study operates by a process different from the pursuit of normative objectives.
- "systematization": Political science research should produce a body of systematic information; theories and generalizations could be based on sound inferences from testable data.
- "pure science": Political science research should operate in a value free manner, that is, independently of any possible subsequent use of scientific knowledge to address perceived social problems.

Robert Dahl (1992) traced the origins of this approach to the 1920s and to the work of Charles Merriman and the so-called Chicago School of Harold Lasswell, Gabriel Almond, V. O. Key, and David Truman. By the mid-1960s, one member of this school—Almond (1966)—was proclaiming "a new paradigm" in political science (p. 875). Almond described this paradigm as having three components: (1) a "statistical approach" geared toward "test[ing] hypotheses" that would generate (2) "probability" statements and (3) a study of the interaction of actors and units within larger political "systems" (p. 876). As is clear in Almond's language, this new behavioral approach was using highly specialized tools and methods drawn from such fields as math, statistics, economics, and psychology. Indeed, Almond pointed out that graduate study in political science was becoming increasingly focused on training students in the tools of "the scientific revolution"—tools that were turning political science in the direction of survey research, statistical sampling, and team-based and grant-funded quantitative research. During the post-World War II behavioralist period, publications in the *American Political Science Review (APSR)* became increasingly oriented toward statistical analyses of public opinion and behavior, especially in the subfields of U.S. politics and comparative politics (Sigelman, 2006). The new focus on studying that which could be precisely and narrowly operationalized seemed worlds removed from the one in which an APSA president could proclaim, as Woodrow Wilson had, his distaste for the term *political science* and his hope for a field of politics characterized by a "Shakespearean range."

A Case Study of Behavioralism: Herbert McClosky's "Consensus and Ideology in American Politics"

Herbert McClosky's "Consensus and Ideology in American Politics," published in the *APSR* in 1964, can serve as a case study for examining more closely the salient features of the behavioralist approach. As the title

of his article suggested, McClosky was interested in the extent to which consensus, or broad agreement, on political values existed in the United States. Although he opened his article with a brief overview of Tocquevillian comments on democratic culture and customs, McClosky framed his analysis around the investigation of specific hypotheses relating to the attitudes of political actors, in this case, actors grouped into two subunits of the U.S. population. McClosky hypothesized that the U.S. public was not uniform in its political views, that it was more supportive of democracy in the abstract than in particular cases, and that political elites (those whom he called *influentials*) were more supportive of democracy than non-elites were.

McClosky (1964) divided the U.S. population into two groups: the influentials and the general electorate. The influentials were individuals who had been delegates or alternates at the major party conventions in 1956, and the general electorate was simply the population at large. McClosky used survey research to measure the attitudes of both groups. With respect to the influentials, a sample of more than 3,000 members of the delegates and alternates at the Democratic and Republican conventions was surveyed. With respect to the general population, McClosky used a national sample of 1,500 adults. Both groups were surveyed on a variety of questions or items, and responses to the items served as “indicators” of “opinions or attitudes” about democratic values (p. 364). If a subunit manifested 75% or higher levels of agreement on an item, consensus was said to be demonstrated.

McClosky (1964) found greater degrees of consensus for democratic procedures among influentials than among the public at large. For example, his surveys contained 12 items to measure support for the “rules of the game” (procedural democracy). These items included statements that respondents were asked to register agreement or disagreement with and consisted of statements about whether a citizen could be justified in acting outside the law, whether majorities had an obligation to respect minorities, whether the means were as important as the ends in the pursuit of political outcomes, whether the use of force was ever justified as a political strategy, and whether voting rights should be expansive or curtailed. Survey results demonstrated, McClosky reported, that influentials expressed consensus on most of the 12 items, whereas the general electorate expressed consensus on none of the 12 items.

McClosky (1964) proceeded to report that, while both influentials and the general population exhibited broader support for freedom of speech when asked about this freedom in the abstract than when asked about freedom of speech for specific unpopular groups, influentials were more supportive than the general population of free speech for unpopular groups. McClosky concluded that one might be led to believe that citizens of the United States had reached consensus on the importance of freedom of speech until one looked at the noninfluentials’ responses to items

involving the application of the principle to particular cases, incidents, and people. For example, support for the rights of Communists, of persons accused of treason, and of convicted criminals was higher among the influentials than among the general population.

Furthermore, McClosky (1964) reported greater consensus among influentials on the importance of the democratic value of freedom than on the democratic value of equality. In fact, McClosky reported the absence of consensus among both influentials and the general electorate on the matter of whether all people were equal, as well as on questions relating to whether all people should be accorded equality. McClosky’s surveys included indicators to measure support for political, social, and economic equality, and his results suggested an absence of consensus among both influentials and the general electorate relating to all three types of equality. In other words, on statements relating to whether most people can make responsible decisions in governing themselves (political equality), whether different ethnic groups are equal (social equality), or whether all people have an equal claim to have a good job and a decent home (economic equality), consensus was absent.

McClosky (1964) also sought to measure what he understood as ideological clarity and the ability to identify oneself accurately along ideological lines. In evaluating survey participants in terms of their responses to particular statements relating to liberal versus conservative issues and their adoption of ideological markers (liberal vs. conservative), he found that influentials were more accurate than the general population in naming themselves as liberals or conservatives and in identifying a position as liberal or conservative.

McClosky (1964) closed his article with six summarizing generalizations. First, elites (influentials) were different from non-elites in terms of a greater elite support for democratic processes and a more complete understanding of political ideology. Second, a comparison of the education and economic circumstances of the two groups suggested possible (and testable) reasons for the differences in attitudes demarcating the two groups. Third, the level of support for democracy among U.S. elites was problematic on some issues (e.g., equality). Fourth, in spite of problematic levels of attitudinal support for democratic values, the U.S. system of Republican–Democratic politics appeared stable, a result, in part, of the nonparticipation of non-democracy-supporting non-elites. In short, democracy, McClosky stated, is sometimes “saved” by the nonparticipation of uninformed segments of the demos (p. 376). Fifth, classic accounts of democracy are inaccurate when claiming that the acceptance of democratic ideas is essential for the survival of democracy. Sixth, although McClosky advised political scientists against becoming sanguine about the lack of support for democratic processes among the population at large, he shared his hope for a wider disbursement of democratic values among segments of the U.S. population

8 • GENERAL APPROACHES IN POLITICAL SCIENCE

as the country continued to promote educational and scientific advancements.

Students of political science can observe key elements of behavioralism in McClosky's work. First, behavior was understood by behavioralists like McClosky broadly enough to encompass opinions and attitudes. Second, it is evident that the turning of the discipline toward the study of the behavior of actors is regarded by behavioralists to be deeply revealing of that which was hidden as long as political science held to traditionalism's tenacious insistence on studying institutions. Behavioralism in the hands of political scientists such as McClosky had accomplished something no less remarkable than to reveal—and prove empirically—the flaws in classic, long-standing accounts of why and how democracies work. Third, behavioralists such as McClosky believed that they had succeeded in demonstrating that big questions such as the ones Wilson wanted political science to address were most reliably answered when turned into narrow, specialized, operationalizable questions and variables. After all, what could be a bigger, more Shakespearean question than the one McClosky had addressed? Yet, only by defining consensus in a narrow, testable way, for example, could McClosky study the question of democratic consensus in such a precise and careful manner. Fourth, behavioralists such as McClosky were not opposed to theoretical generalizations, but they believed that such generalizations were most appropriately developed out of concrete, empirical results; moreover, such generalizations could be used to generate new empirically testable questions. In the process of empirically measuring and testing, however, one was not to allow biases or normative presumptions (e.g., about the goodness of citizens of the United States or of U.S. democracy) to distort one's observations. Finally, the value-free political science of behavioralists such as McClosky tended to produce conclusions that left unchallenged the fundamental structures of the U.S. status quo. As Ricci (1984), Dryzek (2006), and Susser (1992) have noted, behavioralists saw their science as value free but, perhaps ironically, often tended to produce results that fit comfortably with normative assumptions regarding the fundamental soundness of the U.S. political system's ability to address progressively any problems that political science might bring into the open. Indeed, it might even turn out to be the case that what looked like a defect (the apathy of the uninformed) was discovered by means of behavioralism to be an asset.

Postbehavioralism

Definition and Overview

Postbehavioralism is an approach that emphasizes (a) that political science research should be meaningful, that is, that it should address urgent political problems; (b) that

science and values are inextricably connected; and (c) that political science should not seek to model itself on the strict application of scientific methods used in the natural sciences whereby research is driven exclusively by that which can be reduced to narrowly defined questions testable by the most rigorous, most specialized scientific procedures presently available. Postbehavioralists reacted against what they interpreted as behavioralism's excessive reliance on the purity of scientific precision at the expense of "relevance." While many postbehavioralists upheld the value of empirical and statistically oriented research, they tended to argue that behavioralism had overreached in emphasizing a strict adherence to narrow scientific procedures and that behavioralism's proclaimed value-free approach in actuality veiled a normative endorsement of the status quo and was thus both normative and conservative.

A number of postbehavioralist critics of behavioralism, including Peter Bacharach, Christian Bay, Hans Morgenthau, and Theodore Lowi, would join the Caucus for a New Political Science, organized in 1967 (Dryzek, 2006). The caucus continues to conceptualize political science as best carried out when political scientists integrate their identities as community members with their identities as scholars and thus craft research agendas in response to political needs. Political science should be steeped in everyday life and its concerns, not isolated from it as an esoteric, specialized, value-free science, according to Caucus statements (*New Political Science: The Journal*, n.d.).

In 1969, David Easton stated that postbehavioralism was proving to be a transformative force in the discipline. Easton discerned postbehavioralism's presence on two levels: first, postbehavioralism was identifiable as a collection of individual political scientists who shared a growing dissatisfaction with behavioralism's implications, and, second, postbehavioralism was manifested as a new intellectual outlook or approach that could guide research. In his presidential address to APSA, Easton delineated what he called a "distillation" of postbehavioralism's defining elements (p. 1052). Easton described postbehavioralism as a demand for relevance, as forward-looking, as application oriented, and as premised on the belief that it was nothing short of unethical for political scientists to remove themselves from the arena of deliberation and action when confronted with and surrounded by political problems. Easton made multiple references to the Vietnam War, to the threat of nuclear escalation, and to the struggles of the civil rights movement, and he noted that postbehavioralism was an indictment of behavioralism's irrelevance in finding solutions to such problems. Indeed, Easton pointed out that, from a postbehavioralist perspective, behavioralism could be charged with failing even to see such problems, a charge that must have sounded particularly strange to students of McClosky, schooled as they were in regarding influentials or elites as more adept at identifying and understanding political issues than were

members of the general electorate. Easton used the metaphor of blinders to describe what had overtaken a discipline that could not see the obvious, pressing issues of society even while it could describe in copious detail the merits of operationalization, hypothesis formulation, statistical analysis, verification, and falsification. Why, Easton asked, in an era of behavioralism (i.e., 1958–1968), had the *APSR* had only four articles on racial disturbances, only two articles on the practice of civil disobedience, only one article on problems of poverty, and only three articles on urban disorder?

Easton (1969) went on to explain that postbehavioralism's critique of behavioralism was deeply grounded in an understanding of science at odds with that embraced by behavioralism. For postbehavioralists, science was unavoidably based on normative assumptions; thus, according to postbehavioralists, a "value-free" political science (the kind of political science advanced by behavioralists) was not possible. Indeed, postbehavioralists asserted that to proclaim value neutrality was itself a normative stance (i.e., an assertion that a so-called value-free stance was better than its opposite). Postbehavioralism faulted behavioralism for not having acknowledged—and thus not having scrutinized—its own normative foundations and the ways in which those foundations shaped the direction of its research agenda. However, insofar as postbehavioralism was not a rejection of an empirically based science per se, Easton hoped that postbehavioralism could elucidate behavioralism's logic and correct its lack of self-awareness regarding its own assumptions rather than become a repudiation of the gains made in political science's shift away from the early and less scientifically oriented methods of traditionalism. In later years, some scholars would come to regard postbehavioralism's legacy as opening up possibilities of a more "eclectic" application of research methods to the study of political phenomena (Lane, 1990, p. 927).

A Case Study of Postbehavioralism: The Perestroika Protest in Political Science

In December 2000, *PS: Political Science and Politics* published "Voices: An Open Letter to the APSA Leadership and Members." The letter, signed by more than 200 political scientists, had been circulated by someone referring to himself or herself as "Mr. Perestroika." Echoing postbehavioralist concerns from decades earlier, the Perestroika protest letter charged APSA and *APSR* with having a disciplinary obsession with quantitative methodology at the expense of meaningful subject matter. Its narrow methodological focus, the letter argued, had rendered APSA and its premier journal remote from the actual world of scholarly work undertaken by most political scientists. The letter called for increased openness in APSA (e.g., in elections to APSA governing bodies and to the APSA editorial board), the inclusion of a broader range of articles in *APSR*, public

disclosure of survey results that could demonstrate widespread dissatisfaction with the discipline's direction, and greater openness to critical voices in the discipline. Noting that they had not organized themselves into an actual caucus or subunit within APSA, the Perestroika letter signees, nonetheless, claimed to speak for a broad segment of political scientists ("Voices," 2000).

Perestroika supporter Gregory Kasza expanded on the concerns expressed in the initial letter in "Perestroika: For an Ecumenical Science of Politics" (2001). One can see in Kasza's elaboration of the Perestroika protest six major points illustrative of postbehavioralism. First, it was claimed that U.S. political science had been distorted by the dominance within the discipline of highly specialized quantitative research approaches; because of this dominance, Kasza asserted, political scientists seeking to produce scholarly works using qualitative approaches were being marginalized. Second, Kasza argued that the marginalization of nonquantitative approaches constituted a breach of academic freedom. Political scientists, he contended, were being pressured to mold their substantive interests to fit the contours of rigid methodologies and frameworks; he mentioned an anonymous graduate student who had been warned that she would fail as a political scientist if she did not make her dissertation conform to rational choice strictures. Third, in allowing a narrow understanding of science to become dominant within the discipline, political science was undercutting its ability to produce sound scholarship. Indeed, Kasza went so far as to assert that a Perestroika movement could save the discipline from producing subpar scholarship. Fourth, Kasza made the quintessentially postbehavioral call for a political science that was more "relevant" in addressing substantive political concerns. Fifth, Kasza suggested that, in seeking to become as sophisticated a science as possible, political science had actually become something of an adventure in fiction. Kasza charged that scientifically oriented political scientists were, in all too many cases, operationalizing human motives, desires, and choices in such narrow terms (in order to be rigorous) as to render their subjects caricatures.

Finally, Kasza (2001) offered an alternative, "ecumenical" approach. Ecumenism, he explained, would be defined by three elements. First, an ecumenical political science would select problems for analysis and then make decisions about which research approaches would best address the problem, rather than adopting a research approach and defining problems to fit the requirements of the research approach. Second, an ecumenical political science would be explicit in its acceptance of a plurality of methods or approaches. Specialized quantitative methodologies would coexist with qualitative methodologies in an open and expansive political science; for example, graduate programs would reintegrate political philosophy and policy studies into their core areas in a Perestroika-driven discipline. Third, an ecumenical political science would value interdisciplinary study. Kasza urged political scientists to rethink graduate training and, specifically, to institute

dual-degree graduate programs. Political science graduate students should be encouraged to earn master's degrees in alternative and diverse fields, fields encompassing the humanities as well as hard sciences.

Conclusion

In calling for interdisciplinary collaboration, Kasza (2001) was aware that he and other Perestroika supporters were challenging political science to regain something from its earlier orientation. Indeed, in the postbehavioral Perestroika protest, one can recognize remnants of traditionalism. One is reminded of the cross-disciplinary approach of Goodnow when reading recent demands for interdisciplinary breadth in graduate training. At the same time, one can observe in postbehavioralism a parallelism linking the demand to study real people (rather than excessively narrowly operationalized "actors" described by behaviorists) with behavioralism's impatience with traditionalism's earlier preference for studying institutions rather than people. Neither the Perestroika protesters nor other advocates of postbehavioralism purged political science of behavioralism. In fact, at present, one can find all three approaches in political science. One might conclude from a study of the history of traditionalism, behavioralism, and postbehavioralism that political science, as a discipline, has been characterized not as much by complete breaks with preexisting research approaches as by periodic shifts and rearrangements of research emphases (Dryzek, 2006).

References and Further Readings

- Addams, J. (1906). The operation of the Illinois Child Labor Law. *Annals of the American Academy of Political and Social Science*, 27, 69–72.
- Almond, G. A. (1966). Political theory and political science. *American Political Science Review*, 60, 869–879.
- Bay, C. (1992). Politics and pseudopolitics: A critical evaluation of some behavioral literature. In B. Susser (Ed.), *Approaches to the study of politics* (pp. 51–75). New York: Macmillan.
- Bryce, J. (1909). The relations of political science to history and to practice: Presidential address, fifth annual meeting of the American Political Science Association. *American Political Science Review*, 3, 1–19.
- Dahl, R. A. (1992). The behavioral approach in political science: Epitaph for a monument to a successful protest. In B. Susser (Ed.), *Approaches to the study of politics* (pp. 27–46). New York: Macmillan.
- Dryzek, J. S. (2006). Revolutions without enemies: Key transformations in political science. *American Political Science Review*, 100, 487–492.
- Dryzek, J. S., & Leonard, S. T. (1988). History and discipline in political science. *American Political Science Review*, 82, 1245–1260.
- Easton, D. (1969). The new revolution in political science. *American Political Science Review*, 63, 1051–1061.
- Easton, D. (1971). *The political system: An inquiry into the state of political science* (2nd ed.). Chicago: University of Chicago Press.
- Easton, D. (1992). The current meaning of behavioralism. In B. Susser (Ed.), *Approaches to the study of politics* (pp. 47–48). New York: Macmillan.
- Farr, J. (1990). Francis Lieber and the interpretation of American political science. *Journal of Politics*, 52, 1027–1049.
- Fried, A. (2006). The forgotten Lindsey Rogers and the development of American political science. *American Political Science Review*, 100, 555–561.
- Goodnow, F. (1904). The work of the American Political Science Association. *Proceedings of the American Political Science Association*, 1, 35–46.
- Gunnell, J. G. (2006). The founding of the American Political Science Association: Discipline, profession, political theory, and politics. *American Political Science Review*, 100, 479–486.
- Isaak, A. C. (1985). *Scope and methods of political science: An introduction to the methodology of political inquiry*. Pacific Grove, CA: Brooks/Cole.
- Kasza, G. (2001). Forum: Perestroika: For an ecumenical science of politics. *PS: Political Science and Politics*, 34, 597–600.
- Lane, R. (1990). Concrete theory: An emerging political method. *American Political Science Review*, 84, 927–940.
- Lowell, A. L. (1910). The physiology of politics: Presidential address, sixth annual meeting of the American Political Science Association. *American Political Science Association*, 4, 1–15.
- Macridis, R. C. (1992). Major characteristics of the traditional approach. In B. Susser (Ed.), *Approaches to the study of politics* (pp. 16–26). New York: Macmillan.
- McClosky, H. (1964). Consensus and ideology in American politics. *American Political Science Review*, 58, 361–382.
- Melanson, P. H. (1974). Review: The presentation of professional self: Post behavioral currents. *Polity*, 7, 221–233.
- New Political Science: The Journal*. (n.d.). Retrieved March 21, 2006, from <http://www.apsanet.org/~new/journal.htm>
- Ricci, D. M. (1984). *The tragedy of political science: Politics, scholarship, and democracy*. New Haven, CT: Yale University Press.
- Richberg, D. R. (1913). Legislative reference bureaus for political parties. *Proceedings of the American Political Science Association*, 10, 222–233.
- Ross, D. (1991). *The origins of American social science*. New York: Cambridge University Press.
- Shaw, A. (1907). Presidential address: Third annual meeting of the American Political Science Association. *American Political Science Review*, 1, 177–186.
- Sigelman, L. (2006). The coevolution of American political science and the American Political Science Review. *American Political Science Review*, 100, 463–478.
- Susser, B. (1992). The behavioral ideology: A review and a retrospect. In B. Susser (Ed.), *Approaches to the study of politics* (pp. 76–100). New York: Macmillan.
- Ubertaccio, P. N., & Cook, B. J. (2006). Wilson's failure: Roots of contention about the meaning of a science of politics. *American Political Science Review*, 100, 573–578.
- Voices: An open letter to the APSA leadership and members. (2000). *PS: Political Science and Politics*, 33, 735–741.
- Wilson, W. (1911). The law and the facts: Presidential address, seventh annual meeting of the American Political Science Association. *American Political Science Review*, 5, 1–11.

2

POSTMODERNISM

ROBERT SWAN

Western Oregon University

“Postmodernism,” writes criminologist John Crank (2003), “is a body of philosophy, methodology, and critical review of contemporary society that encompasses a variety of standpoints” (p. 153). Although we will revisit this simple description of postmodernism in some detail below, it is not uncommon that when first encountering this (or similar) encapsulations of postmodernism, many students of political theory are left scratching their heads. This is not necessarily the fault of the student. In fact, scholars, too, are left scratching their heads (sometimes angrily) over the dilemma of postmodernism and its “questionable” application to “real life.” Whether postmodernism and postmodern theories are applicable to real life is a debate, essentially, about the nature of reality and the value of some types of knowledge over others. This chapter of *21st Century Political Science: A Reference Handbook* intends to plunge the student directly into this debate. Drawing inspiration from famous postmodernist Jean-François Lyotard, this chapter intends to expose readers to knowledge that will both enhance their knowledge base and change the way they acquire and process knowledge in the future.

As a serious student of political science, the reader is likely referencing this handbook in order to answer specific questions about postmodernism. The bad news is that philosophical postmodernism rejects absolute answers to almost any question. The good news, on the other hand, is

that exposure to postmodern thought (and its application to empirical research) will broaden, and thus enhance, the reader’s knowledge of the world. Simply, while postmodernism may reject dominant narratives (i.e., “official” answers), it offers a great deal of insight into many social worlds that have gone largely unexamined. This handbook then will increase the reader’s level of sophistication regarding the “what is” and the “what ought to be” as conceptualized by postmodernist scholars. These are not unimportant questions in political science, and an enhanced knowledge of how postmodernism has influenced the way these realities are constructed will enhance the reader’s ability to think critically about political and other social phenomena.

This chapter summarizes the broad topic of postmodernism and distills it into its essential elements. First, it reviews the literature regarding what postmodernism is, in both a temporal and a theoretical sense. This review includes some of the common elements found in postmodern thought and writing, as well as some of the key differences among postmodern thinkers. Next, the chapter discusses the application of postmodern thought to empirical research. The large body of contemporary research influenced by postmodernism cannot be reviewed here. However, empirical work deriving from social construction theory is an area in which postmodernism has been most influential, and this chapter discusses public perceptions of crime and politics and the application of social construction

theory to our understanding of institutional life. Next, it discusses the policy implications offered by postmodern thought and research and, in particular, policy implications of the key postmodern assumptions that reality is entirely subjective and that there is a dark side to our existence in postmodern bureaucratic systems. Indeed, from a postmodernist perspective, dominant narratives and discourse blot out a myriad of politically and socially important experiences, perspectives, and voices of both individuals (e.g., prison inmates, radical political actors, sex offenders) and entire groups of people (e.g., undocumented immigrants, women, and ethnic and racial minority groups). Simply, postmodernists draw our attention to concerns that we would otherwise not consider. The chapter also examines future directions for postmodern theory in relation to the mass media and ends with a summary and a conclusion regarding the importance of understanding postmodernity as both a temporal and a theoretical frame of reference.

Postmodern Theory

[The postmodern world] is less a world of facts and figures and more a world of story and performance.

William Bergquist (1993, p. 23)

Although there is little consensus on its origins, the concept of postmodernism began to be used in the late 19th century and has been embraced by a wide variety of fields, including architecture, visual art, literature, philosophy, political science, sociology, fashion, and many others. In the postmodern era, two primary philosophical positions compete for dominance: objectivism and constructivism. Postmodernists attribute *objectivism*—the notion that we can objectively determine reality, discover universal truths, and make sound decisions based on our findings—to modernity. For postmodernists, objectivism is an illusion. Rather, *constructivism*—the notion that individuals live in unique realities and construct these realities based on the situation in which they find themselves (e.g., race, class, gender, social context)—provides a more comprehensive intellectual platform from which we can understand the nature of reality. For postmodernists, there is no universal truth but rather billions of individual truths.

From a philosophical perspective, and as it directly applies to the social sciences, many scholars highlight French philosopher Jean-François Lyotard for his articulation and popularization of postmodernism. Before his expression of postmodernism is discussed, it is useful to briefly discuss the general controversies surrounding the postmodern movement and some of its essential intellectual components and applications.

Many positivist scholars turn away from postmodernist frameworks because of postmodernism's rejection of

positivist assertions regarding scientific objectivity (and positivism's strivings toward that ideal), its lack of consistency as a theoretical framework, its lack of linguistic and empirical clarity, and the general disagreement in postmodernist literature on what postmodernism actually is. Simply, the concept of postmodernism is an often disputed and nebulous concept incorporating a number of temporal, spatial, and theoretical elements related to our understanding of multiple (if not countless) subjective realities. In fact, the concept is often debated, sometimes vigorously, within the postmodern literature itself. Additionally, the term *post modernism* is a broadly defined concept that incorporates a number of "strains" and subdisciplines, most of which are based on normative and subjective evaluations of past and present "realities"—all of which are individually subjective and subjectively interpreted. Simply, postmodern theories—in all their many forms—generally suggest an unmanageable subjectivity that could be problematic for our understanding of political phenomena and, indeed, for the maintenance of a cohesive social system.

That being said, postmodernism is an intriguing and potentially useful philosophical approach to understanding current phenomena such as the impact of globalization on social relationships, the behavior of political actors and institutions, and the origins and behavior of criminal justice institutions and policies, to name only a few.

For the sake of clarity, it is useful to briefly unpack the notion of postmodernism in its temporal, spatial, theoretical, and conceptual senses, beginning with modernity. "Modernity," asserts new-modernist Anthony Giddens (1990), "refers to modes of life and organisation which emerged in Europe from about the seventeenth century onwards and which subsequently became more or less worldwide in their influence" (p. 1). However, the term *modernity*, asserts Philip Selznick (1992), "is not a synonym for the 'contemporary' or 'present day'" (p. 4). If that were the case, the term *modernity* would be meaningless in the sense that people living in every historical epoch would consider themselves to be living in a modern period. Modernity, perhaps best marked by the Enlightenment and the advent of rational, scientific thought, global exchange, mass media, and mass production, served to break up previous forms of traditional, precapitalist and premodern life (e.g., small, isolated villages and communal living). Selznick cites four reasons for the transformation from the premodern to the modern society: (1) "structural differentiation" where there was none before; (2) "secularization" and religion's loss of power over governance systems; (3) "atomization" and the "weakening of social ties," which facilitate the likelihood that the number of shared experiences will diminish and, when present, will be unlikely to continue to have a bonding effect; and (4) increasingly coercive "rational coordination" through "contract and bureaucracy" (pp. 4–5), which has facilitated experiential fragmentation among individuals working within increasingly large organizations, as well as among those who must be served by them.

Like the terms *modernity* and *modernism*, the term *post modernism* generally indicates both a historical period and an intellectual position in any given field. Simply, the post-modern historical period followed modernity and is closely associated with rapid advancements in technology, increased surveillance, rapidly accelerating globalization, atomization, increasing social disconnection between people and places, and the increasingly rationalized, bureaucratic state. In theory, the postmodern *condition* (as opposed to temporal and spatial notions of postmodernity), according to Giddens, is, in part, related to society's generalized angst and confusion about the world and individuals' perception that they are unable to control, or even make sense of, their own destinies. In short, argues Jean-François Lyotard, modernity's "grand narrative"—the dominant, linear stories that we had constructed during the modern period about our history and (predictable) prospects for the future—have evaporated in the postmodern period. The evaporation of the grand narrative—or "the murder of reality," as Jean Baudrillard (1996, p. xi) refers to it—has led to a great deal of social angst and uncertainty.

Interpreting and applying postmodern thinking in the social sciences is no easy chore. However, one way to begin thinking about its application in the social sciences is to think about the possibility that the dominant narratives (e.g., "sex offenders cannot be cured" or "Iraq, Iran, and North Korea are the axis of evil") may not be accurate, and, it is important to note, are not the only important stories that need to be told in order to maintain social cohesion. Jonathan Simon (1997), for example, simply argues that postmodern thought revolves around concern over "what one thinks has changed in the present that requires breaking the useful interpretive frames that have been [traditionally] associated with modernity" (p. 171). One good example of applied postmodernism can be found in critical criminology. Critical criminologists, drawing from the assumptions of postmodernism, assert that hegemonic normative values, inextricably and historically linked to capitalism, are doing a great deal of damage to marginalized individuals, or rather, those individuals who do not conform to the grand narrative (e.g., racial and ethnic minorities living in poverty). Postmodernism in criminology then comprises theories that attempt to expand our conception of what is or is not criminal and what is or is not just. Thus, it should be no surprise that some criminologists are skeptical of penal strategies that attempt to predict or "solve" crime based on rigid and narrowly defined categories of *crime* and *criminal*—especially those that appear to sacrifice individualized justice or the greater collective good.

Postmodernism runs a similar course through the field of political science. Postmodernism in political science, as in the other social sciences, typically appears as forms of thought or of empirical research that question dominant narratives and seek out alternative voices and perspectives in order to enrich our political discourse through the inclusion of previously marginalized people. In addition to

recognizing the importance of subjective, individual realities, postmodern theories highlight the role of conflict in social life—an important underlying consideration for many students and scholars who study complex societies and their institutions. For example, John David Farmer and others argue that in understanding modern organizational forms and purposes, postmodernism can help us understand how (and why) the instrumentalism of modern bureaucratic structures limits the ability of human beings to self-actualize. This postmodern focus is supported by a number of scholars in multiple and varied fields who, although not necessarily postmodernist, continue to emphasize the negative effect that instrumentally oriented organizations may have on society and on those who work (or who are imprisoned) within them.

For example, in articulating a defense of postmodernism in public administration, Farmer (1997) asserts that postmodernism is a response to the dominant and oppressive narratives that have defined the modern period. Primarily, these narratives have had a negative impact on women, ethnic and racial minorities, sexual minorities, and the poor. In particular, Farmer is critical of the rigid, hierarchical form that bureaucratic structures typically take and the lack of focus on what he calls the bureaucratic "in-between." Farmer's focus on the bureaucratic in-between revolves around the idea that the study of bureaucracy should begin to focus more on the individual experiences of bureaucratic workers and less on bureaucratic technology and efficiency. In this way, we can run less oppressive—and thus more effective—bureaucratic institutions. This concern is somewhat derivative from the Hegelian notion that "man" will fight to the death in order to be recognized as something other than a slave and recognizes that social conflict results from the "enslavement" of marginalized peoples. As Farmer, Max Weber, and others see it, bureaucrats should be included in our understanding of oppressed peoples, primarily because they are enslaved and marginalized in a postmodern society. In suggesting reforms, then, postmodernists attempt to eliminate social conflict by allowing individuals to be recognized, considered, and treated as unique entities with unique perspectives and needs rather than as efficient (or inefficient) organizational instruments to be thought of only in aggregate, actuarial, or economic terms.

However, Frank de Zwart (2002) writes, "Postmodernists confuse wrongs of bureaucracy with arguments against modern science and then propagate relativism to clear up the muddle they created" (p. 482). In many ways, de Zwart is correct. Because postmodernism does not seek to obtain generalized knowledge, many argue that an entirely postmodern focus in the social sciences may leave us "empty-handed" in terms of usable facts. Simply, how does a society create sound policy based on anecdotal evidence? Obviously we cannot. But gaining an understanding of marginalized individuals and groups can help us expand the grand narrative to be more

inclusive and, perhaps, less destructive to the social life of those who are marginalized. This, some postmodernists argue, will eventually benefit our entire society. For example, some postmodern approaches can help us better understand how and why prisons evolved, how the courts have responded to a variety of individuals and social conditions, and how and why the mass media generate distorted, yet influential, images of some phenomena but ignore others altogether.

To give students a working knowledge of a fairly vast and complex body of work, the next section identifies and delves into the work and intellectual assumptions of some of postmodernism's more widely recognized thinkers. Further readings suggested at the end of this chapter provide more depth.

Postmodern Thinkers: Lyotard, Derrida, Foucault, Baudrillard, and Giddens

Postmodernists (broadly speaking) are not interested in building theory in a traditional sense. That is, they reject positivism and are not interested in building theory through what Thomas Kuhn (1962) has referred to as *normal science*. Rather, postmodernists are interested in studying the anecdotal (e.g., individual experiences and perceptions, media portrayals of phenomena, and language construction and usage) and the intellectual implications that these phenomena have for social life—especially among marginalized populations. It is important to note that postmodernists are quite willing to change their minds, about the conclusions they come to initially and recognize that knowledge is socially constructed and, thus, fallible. The following five thinkers are generally recognized as some of social science's most influential postmodern thinkers. However, and as discussed previously, postmodernism is a somewhat broad and nebulous concept, and the term *postmodernist* is rarely attributed to any of these philosophers (with the exception of Lyotard). That being said, all the thinkers discussed below are postmodern in the sense that they all reject positivist methodologies as the only valid form of knowledge acquisition. Similarly, they all reject modernity's grand narrative and explicitly or implicitly offer support to Lyotard's contention that the collapse of the grand narrative marks a new historical epoch.

Jean-François Lyotard

Our working hypothesis is that the status of knowledge is altered as societies enter what is known as the postindustrial age and cultures enter what is known as the postmodern age.

Jean François Lyotard (1979, p. 3)

French philosopher Jean-François Lyotard was concerned with articulating a coherent conceptualization of postmodernism and is widely regarded as having been successful in doing so. In *The Postmodern Condition*, Lyotard discusses the changing nature and acquisition of knowledge in the postmodern period, due, in large part, to rapidly changing technology and the resulting social transformation. Of interest in this work is Lyotard's discussion on the changing nature of knowledge. In a postmodern, technological age requiring more and more technical knowledge, knowledge for knowledge's sake is becoming less important to social survival and is being replaced by knowledge that may be sold to purchasers who seek to put it to work. In the social sciences, we see this transformation very clearly as knowledge becomes distinguished by its applied or theoretical nature. In addition to our ranking the natural or hard sciences over social sciences because of utility concerns, what we see is a trend toward valuing applied social science research far more than we value theoretical or philosophical social science research—a trend that many scholars see as disastrous for the acquisition of future knowledge. Implicit in Lyotard's discussion in *The Postmodern Condition* is the idea that those who possess the skills and knowledge needed to produce applied research will garner most of the social, political, and economic power, essentially leaving those who pursue knowledge for knowledge's sake outside the circle of power. In an ironic twist, then—irony is also a key feature of postmodern thinking—as we move away from the acquisition of broad-based, theory-driven knowledge toward increasingly specialized and applied descriptive and technical knowledge, we also decrease our ability to create generalizable knowledge in the future.

In other work, Lyotard has focused on the role of language in our acquisition and understanding of the world around us. In particular, Lyotard has focused on the collapse of the grand narrative, or rather, the *metanarrative*, due to social atomization (i.e., individual isolation) and the resulting rise of *micronarratives*. Micronarratives, according to Lyotard, represent the rejection of grand narratives at the individual level and the acceptance and integration of knowledge by an individual only as it relates to the individual's particular circumstance (i.e., ideology, race, class, gender, experiential realities). As discussed below, understanding the micronarrative is a key focus in constructivist research and is extraordinarily important to, for example, our understanding of how the media influence public perceptions of any given phenomenon.

Jacques Derrida

The history of writing should turn back toward the origin of historicity. A science of the possibility of science? A science of science which would no longer have the form of logic but that

of grammatics? A history of the possibility of history which would no longer be an archaeology, a philosophy of history or a history of philosophy?

Jacques Derrida (1967/1997, pp. 27–28)

As is the case with all the postmodernists discussed here, French philosopher Jacques Derrida questioned the notion that there is an objective truth. Rather, social reality is highly subjective and prone to abuses by powerful, self-serving elites who use their power to help society construct dominant narratives about reality. Unfortunately, these grand narratives are used to oppress or enslave social minorities, non-elite workers, women, and the poor, among others. Throughout history we can see that some “truths” were simply not that accurate at all. For example, the medieval notion that animals were as culpable (and thus, as punishable) as the humans who abused them in medieval bestiality cases, or that “quirky” women were witches, simply did not pan out over time. However, these “truths” did serve to facilitate the power interests of elites for a time. The same can be said of the United States’ early and long-lasting use of prisons as places of “rehabilitation,” or the perpetuation of Black slavery based on the accepted “truth” that African American slaves were less human than Whites. This latter, erroneous “truth,” perhaps more than any other social construct associated with American slavery, continues to haunt and create conflict within our society to this day.

According to postmodernists, we continue to live in a world rife with dominant and harmful mythologies designed to serve powerful elites. To cite one example, today, most citizens of the United States mistakenly believe that sex offenders (broadly defined) are incurable. Unfortunately, this “truth” is unsupported by empirical research yet has led to a large number of lifelong penal sanctions (including the possibility of the death penalty) for a large number of people convicted of sex offenses ranging from indecent exposure and simple kidnapping to the most serious types of rape and child molestation. It is interesting that the technologies of control and penal leniency (Foucault, 1977/1995) designed to control sex offenders have now begun to move to other offenders and, ultimately, will likely be used to maintain control over society more generally (see Diana Gordon, 1990). One recent example is a proposal by South Dakota corrections officials to make identities, addresses, and criminal histories available about all offenders, not just sex offenders, via online websites.

Language—and the manner and means in which it is delivered—is an important part of how society constructs reality and an important part of understanding the essential focus of postmodernism. Derrida, like Lyotard and Baudrillard, was concerned with the role of language in

our society and the way language is used to construct reality. In particular, Derrida was concerned with *deconstruction*—an examination of the underlying meaning and foundations of language, text, symbols, and other signs—in order to show that the dominant interpretation and foundational logic were flawed. In other words, multiple interpretations and meanings are possible and the foundation on which dominant interpretations rest is not solid ground but rather nothing more than subjective and biased beliefs.

Derrida, essentially, was criticized by peers and laypeople alike for his philosophical position—often likened to nihilism—that we cannot really know anything. In particular, academics attempting to generate useful, if not generalizable, knowledge vigorously opposed Derrida’s position and philosophy. Indeed, Derrida poses a somewhat serious epistemological problem for social scientists attempting to conduct, analyze, and interpret research and its findings: They all rest on subjective and biased foundations and so cannot really be true. Derrida’s intellectual position on the acquisition of knowledge generated a great deal of controversy precisely because it devalued all acquired knowledge.

Michel Foucault

In my view one shouldn't start with the court as a particular form, and then go on to ask how and on what conditions there could be a people's court; one should start with popular justice, with acts by the people, and go on to ask what place a court could have within this.

Michel Foucault (1980, p. 1)

Foucault’s (1977) understanding of power, influenced to a large degree by Friedrich Nietzsche and Max Weber, is linked, somewhat ironically, to the rise of an enlightened and egalitarian society in the 18th century. That is, argues Foucault, the rise of the egalitarian state in the 18th century also gave rise to distinctly less-than-egalitarian forms of social control. In order to ensure that the rights of “all” were respected, it was necessary to segment and discipline society in order to control, correct, and monitor transgressors. The disciplining and “correction” of transgressors set an example for society and ensured a more disciplined society. Therefore, in implementing egalitarianism, society in fact became more repressive and repressed. However, argues Foucault, these methods of control, correction, and surveillance were not born out of thin air and in fact had been present in very diffuse form for quite some time. Thus, Foucault’s understanding of power is heavily reliant on an understanding of preexisting forms of social control and their systematic linkage and evolution through reform (e.g., penal leniency) in the 18th century. In doing so,

argues Foucault, the egalitarian project merely enabled a more systematically intrusive governance system and repressed society.

In *Discipline and Punish*, Foucault (1977) shows how irregular forms of punishment were consolidated and “reformed” in order to better regulate and economize the state’s use of power in its attempt to transform and control society. The reformation of technologies of power was not so much about the elimination of certain behavior as about modulating social behavior so that it could more efficiently serve the needs of the state. In reforming punishment, then, the point was not so much to eliminate crime or criminals as to ensure that the efficient and economical regulation of crime and criminals served the state’s and, ultimately, society’s needs. Simply, crime and criminals serve the economic and political needs of the state, as well as establish a model, or an example, for the rest of society of how *not* to behave. This new “political economy” was facilitated by the development of *panopticism*. Panopticism, in relation to 18th-century reform, draws on, and integrates, notions of internment, hierarchical differentiations, and inescapable surveillance, much as the “new penology” does today. Thus, Foucault’s post-modern analysis regarding the rise of prisons not only helps us understand the driving force behind mass incarceration in our society but also gives us insights into the driving force behind all our institutions.

Jean Baudrillard

Reality is a bitch. And that is hardly surprising since it is the product of stupidity's fornication with the spirit of calculation—the dregs of the sacred illusion offered up to the jackals of science.

Jean Baudrillard (1996, p. 3)

Like Derrida, French philosopher Jean Baudrillard questioned the notion that there is an objective truth. Rather, social reality is highly subjective and prone to abuses by powerful, self-serving elites who use their power to shape society’s construction of dominant narratives about reality. In particular, Baudrillard was interested in *semiotics*, the study and understanding of how words (*signs*) interrelate. Specifically, semiotics is not necessarily interested in what words themselves mean but rather what they mean in relation to one another. For example, a semiotic approach to understanding an individual’s reality would assert that if an individual thinks about his or her automobile, the individual is actually thinking about those things that are not his or her automobile (e.g., home, spouse, school). This is because in order to construct an image of the automobile, an individual must locate it within a previously constructed web of meaning. This focus is evident in Baudrillard’s work, especially in his analyses of mass media.

In *The Perfect Crime*, Baudrillard (1996) is concerned that we have overburdened ourselves with meaningless,

confusing, referential imagery and positions the age-old philosophical question, “Why is there something rather than nothing?” (p. 2) as a straw-man argument. Baudrillard rebuts his own question with another: “Why is there nothing rather than something?” (p. 2). In answer, he argues that the perfect crime has been committed. In fact, he argues, we have murdered reality—false though it may have been to begin with—by extinguishing the grand illusion. The grand illusion, according to Baudrillard, is a personally coherent set of referential signs and meanings—to include the cherished notion of an *objective* reality and the formal illusion of truth, which we have traditionally used to hide from grim social realities. This “murder,” he argues, has been accomplished through the swamping of society in a sea of virtual and meaningless imagery (e.g., fantasy video games, the trials and tribulations of Paris Hilton, infomercials, Facebook). It is a *perfect* crime because the sea of meaningless images masks the “murder” of society’s mythical image of an objective reality. In sum, society continues to believe, generally, in an objective reality but only on an individual level. Reality, then, is very individualized, and perceptions of reality are unlikely to be shared on a very large scale.

Included in the idea that we are left to our own devices when constructing individual realities is the notion that we construct our individual realities based on imagery (signs and symbols) that were not real to begin with. In *Simulacra and Simulations* (1981), Baudrillard suggests that we are creating “bad” copies from false images—images originally constructed to mask the fact that there was nothing there to begin with. There are a number of rather poignant examples we can use to illustrate this point. The one that may resonate most with students is the notion that men and women seek physical inspiration from the sea of visual imagery depicting perfect bodies in perfect health. Unfortunately, very few of these images are real and are most likely airbrushed or digitally enhanced in some way. Thus, many of us attempt to personally re-create a physical reality that never existed to begin with, and inevitably we “produce” (e.g., through cosmetic surgery, obsessive dieting and exercise, and the conspicuous consumption of name-brand products) incomplete physical “copies.” Needless to say, the results of this “copying” process have led to a great deal of social and personal angst (e.g., the inability to accurately simulate our favorite supermodels, media stars, or sports heroes).

Anthony Giddens

The views I shall develop have their point of origin in what I have elsewhere called a “discontinuist” interpretation of modern social development. By this I mean that modern social institutions are in some respects unique—distinct in form from all types of traditional order.

Anthony Giddens (1990, p. 3)

British sociologist Anthony Giddens is a critic of modernity in the temporal sense and of postmodernism in its theoretical and philosophical sense. However, he is, arguably, a postmodern thinker in the sense that he places a great deal of importance on the study of power, knowledge acquisition (and the influence of power on knowledge acquisition), and the influence of knowledge on people's ability to alter their individual and collective social and material reality. In particular, Giddens confronts the question of how social reality is constructed. In his theory of *structuration*, Giddens's main area of concern is whether social reality is primarily influenced by individuals or by broad social forces.

In *The Consequences of Modernity*, Giddens confronts the issue of whether we are in a postmodern period, as Lyotard and others argue, or whether we are experiencing a type of radicalized modernity in which modernity has simply accelerated its pace. This radicalized modernity is primarily fueled by the disembedding processes of globalization, best characterized by global, cultural homogeneity; the disappearance of tradition; the erosion of place-based community; and the erosion (and shifting nature) of trust in persons, institutions, and abstract systems. In sum, radicalized modernity, although postmodern in a temporal sense, is not postmodern in a theoretical sense, because it is simply a continuation of modernity, not a temporal epoch that will usher in a dramatically different type of social order (such as the transition between premodern and modern societies).

Unfortunately, argues Giddens, riding the "juggernaut of modernity" has a dark side. Primarily because of the increasing lack of trust in abstract systems, institutions, and people, the accelerated pace of modernity is likely to usher in one or more forms of totalitarianism. We can see hints of this today from the expanded use of closed-circuit television on street corners and other highly engineered social environments, developed by governments in response to perceived (or contrived) security concerns. (Jonathon Simon, 1997, discusses this governance style as being very similar to an airport model of governance.) What this means, simply, is that much like Stalin's Soviet Union, internal passports (i.e., one or more forms of valid, government-issued identification) must be carried by individuals so that authorities can assess an individual's identity and background on a moment's notice and prior to letting that individual proceed in peace.

Applications and Empirical Evidence: The Construction of Social Power and Oppression

As Ray Surette (2007) has so eloquently put it, "Reality is a collective hunch" (p. 34). Perhaps no other sentiment exemplifies the empirical focus of scholars who view (and research) social phenomena through a postmodern lens. In other words, social and political reality is entirely subjective. In a democratic system subject to the constraints of people's beliefs about any given phenomenon, this may be

a problem. In fact, many postmodern thinkers challenge the idea that policies influenced by a majority of people who believe the same thing are good policies. In fact, postmodernists tend to think these policies are quite destructive both to marginalized individuals and to society as a whole. Primarily, they argue, this is because people have been subject to, and influenced by, powerful actors who attempt to dominate mediated policy discussions for reasons of power, economic profit, or both. Four areas of research influenced by postmodern thought illustrate these themes.

Socially Constructing Political Reality Through the Nightly News

Scholars conducting news media research are concerned with both *how* news media organizations represent reality and *how* the public makes sense of these depictions, not to mention the influence that both of these phenomena may have on policymakers and on public policy generally. Media scholars have generated a great deal of empirical data regarding the influence of the news on public policy. Generally speaking, the consensus among scholars is that the news, depending on the individual attributes of the viewer, has a more or less significant effect on people's attitudes toward any given phenomenon, which in turn may have an effect on the political agenda and on policy and election outcomes. Drawing from constructivist theory, many news media scholars view the news media (printed, electronic, and "new" media) as a sort of gladiatorial arena in which a variety of competing viewpoints vie for time and space. As media scholar Regina Lawrence (2000) puts it, "[The news] is an arena of problem construction in which struggles to designate and define public problems are waged" (p. 3). The "winners" of these competitions tend to dominate the policy process and, if consistently dominant in the media, win clear electoral, legislative, and policy victories on a fairly regular basis. Unfortunately (from a normative, postmodernist perspective), and with few exceptions, those who tend to dominate media also tend to be social, economic, and political elites.

The Social Construction of Crime Problems

Scholars who study the social construction of crime problems are typically concerned with why the public misperceives the actual incidence and seriousness of crime, how the public has constructed this distorted image, how these images influence crime control policy, and why, in many cases, these distorted images are resistant to actual data. In addition to focusing on news and other media portrayals of crime and criminality, scholars who study the social construction of crime problems are interested in the influence of individual attributes (e.g., age, race, gender, social context) and personal experiences on perceptions of crime as well as the way people think about crime and talk about crime with one another.

*The New Penology: Rabble Management
and Actuarial Risk Assessments*

Simply stated, the new penology is both a critical, postmodern theoretical framework and, in a temporal sense, an emerging set of penal policies related directly to the conditions of postmodernity. As a theory, the new penology is considered a critical or postmodern theory (depending on the epistemological assumptions of any given approach—e.g., *critical theories* are more tightly coupled to Marxist assumptions). Postmodern theories are used to analyze and critique postmodern penal strategies and asymmetrical uses of power, as well as to suggest reforms. The new penology, critical criminologists and penologists argue, is a postmodern penal strategy worthy of critique.

According to Feeley and Simon (1992), the new penology emphasizes penal language that moves away from the traditional focus of criminal law—individualized justice concerns such as due process, punishment, redemption, and rehabilitation—and focuses more on the management of risky aggregates (e.g., actuarial tables and risk assessment instruments designed to detect actual and potential criminal offenders by categorical grouping). Typically, arguments critical of the new penology assert that individualized “justice” cannot be achieved through current policing, adjudication, and incarceration processes due to a widespread institutional failure to recognize any moral dimension (i.e., moral deficits) to criminal offending or an individual’s unique characteristics and circumstances. The recognition of a moral component to criminal offending and the uniqueness of individual offenders have been the traditional focus of British and American transformative penal practices. These practices date back to (at least) U.S. colonial reforms implemented by the Quakers in the 17th and 18th centuries, which emphasized transformative punishment through solitary confinement, individual repentance, and successful reentry. The focus on individual offenders and their moral transformation continued onward through 19th-century Jacksonian-era institutional prison reforms and 20th-century Progressive-era indeterminate sentencing reforms—all of which continued to emphasize both the moral components of offending and the social requirement to transform each individual into a productive, conforming member of society. The moral dimensions of offending, as well as the individual characteristics and circumstances of offenders, it is argued, are now ignored by institutional actors, having been replaced by profit- and budget-maximizing concerns, actuarially based risk assessments, and other aggregate management techniques that favor institutional management interests over individual and community justice. For example, the potential riskiness of convicted sex offenders is determined through the use of actuarial tables and statistically advanced risk assessment instruments. Based on the statistically determined risk level of an individual, traditional justice concerns may be minimized or ignored. This is especially evident in policies that do not factor “needs” into a risk and needs assessment and policies that dictate

the indefinite (or lifelong) civil commitment (detention) of sex offenders even after they have served their prison time.

The new penology, echoes Lisa Miller (2001), is unconcerned with reducing crime through traditional rehabilitation programming or more comprehensive social programs. Rather, argues Miller, the new penology “is aimed at simply managing the harm that crime inflicts” (p. 170). Toward this end, the practices and policies of the new penology emphasize the empowerment of the U.S. culture of control (Garland, 2001; Gordon, 1990) through increased surveillance and actuarially based risk management that, it is hoped, can be used to prevent crime altogether. As summarized by Miller (2001), the new penology, in characterizing all aspects of criminal justice policy, relies only on actuarial precision and efficiency in order to manage and contain risky groups and individuals—what Lynch (1998) has also referred to as a “waste management” model of criminal justice.

Robert Bohm (2006), in tapping into the very essence of the new penology (without actually ever calling it so), argues that a “McDonaldization” of criminal justice has occurred. His thesis revolves around the idea that a hyper-efficient criminal justice system—reminiscent of the somewhat sinister, Weberian depictions of the technical-rational bureaucracy—has evolved to more efficiently handle increasing numbers of offenders and potential offenders. This efficiency, primarily handled by nonhuman technologies—technologies that are becoming increasingly advanced—allows elites to more effectively control society by creating a criminal justice system that is more calculable and predictable, although less just. Bohm ultimately argues that this new rationality has led to a great deal of irrationality in the sense that McDonaldized criminal justice systems do more social harm than good.

Implementation of new penology policies and practices can have unpleasant consequences for everyone but is especially problematic for particular groups and classes of people. For example, using racial criteria (e.g., racial profiling) as a determinant or an indicator of bad character or criminal intent, though legal (with some qualifications), is widely condemned as unfairly stigmatizing an entire group of people. Similarly, John Irwin (1985, p. 2) asserts that the use of police and jails is a form of “rabble” (i.e., people of lower socioeconomic status) management and that prisons have become “warehouses” (Irwin, 2005, p. 2) and receptacles for the “new dangerous classes” (p. 8). Finally, the most vehement critiques—if usually only implicit—surrounding new-penology crime control strategies revolve around the notion that these strategies may destabilize communities and actually create more, rather than less, crime and violence (Clear, 2007), as well as resistance to all formal authority (e.g., the No Snitch Movement in some African American communities).

Understanding Institutional Realities

Research influenced by postmodern assumptions can be both interesting and useful to the successful management of

our political and social institutions. For example, Earl Babbie (2007) has commented that institutional ethnography—an approach developed by Canadian sociologist Dorothy Smith as an approach designed to study the everyday experiences of women—was initially intended to better explain the institutional experiences of “oppressed subjects” (p. 300). Institutional ethnography then incorporates a variety of methodological approaches designed to capture the personal experiences of those who do not often (if ever) get to speak about their experiences. Dorothy Smith (2006) writes, “Institutional ethnography explores actual people’s activities as they coordinate in those forms we call institutions” (p. 14).

From the perspective of Smith and others, institutional ethnography is a postmodern form of critical microsociology that largely draws its inspiration from Marxism and the assumptions of postmodernism. As a standpoint approach to research and action (Harding, 2004)—theory building and activism based on understanding the world through the eyes of oppressed peoples—its effort to link the microsocial experiences of marginalized subjects to macrosocial structures and trends has generally been well regarded.

Policy Implications

In our discussion of postmodernism thus far, many of the policy implications may be self-evident to the reader. On the other hand, conceptualizing theoretical, philosophical, and applied postmodernism is a vast undertaking, and it is useful to briefly summarize some of its key policy implications. However, because postmodernism discourages the belief that we can objectively “know everything” (or anything, in the case of Derrida), that is, knowledge is subjective and dynamic, the following list is not all-inclusive. Rather, it is designed to inspire thinking about a few of the possible policy implications suggested by postmodernism:

- *Language:* Postmodernism’s focus on linguistics underscores the policy problems associated with language and the role of language in the social construction of reality at the micro- and macrolevels. In particular, Baudrillard’s focus on the mass media suggests that in a postmodern age, the mass media may be a highly problematic source of information—an assertion that has been supported by a number of media scholars conducting theory-driven empirical research.
- *Radical modernization:* Giddens’s idea that we live in an age of radical modernization suggests that we will become more and more disconnected from one another on the individual level, while becoming more and more homogenized and integrated on a global level. In sum, radical modernization, driven by radical globalization, will atomize individuals even while allowing us to check the weather in Bangkok

and to “facebook” with “friends” in Australia. The downside to radical modernization is that doubt, as Giddens argues, will become institutionalized on a global scale, and trust in individuals, institutions, and abstract systems will be eviscerated. What this may mean in the future is that we will continue to have more surveillance, less freedom, and an increasingly diminished ability to collectively solve social problems.

- *Knowledge for knowledge’s sake:* As Lyotard suggests, the pursuit of knowledge for knowledge’s sake may be a pursuit of the past. In the postmodern age, knowledge must be practical, applicable, and if at all possible, technically useful. Most important, this type of knowledge has monetary value and is generated in order to be sold. While all this may sound fine, there is a dark side to abandoning the pursuit of knowledge for knowledge’s sake. In particular, we can see the impact of this trend on university campuses everywhere as some departments—such as English, sociology, and history—continue to lose their places as important components of a university education, even among traditionally liberal arts schools.
- *Bureaucratic efficiency versus individualized justice:* Related to many of the concepts previously discussed, another dark side to postmodernity (and a key focus of public administration scholars and criminologists) is the focus on bureaucratic efficiency over individualized justice. In other words, the increasing focus on economic and technical efficiency in all our social institutions may cause a great deal of what Adams and Balfour (1998) have termed “administrative evil.” Simply, administrative evil is the disconnection of means from ends. From the critical perspective of postmodernists, it means that the ends, regardless of how these were conceived, are more important than the means institutions use to achieve their goals. Many political science, public administration, and policing and prison scholars, not to mention critical criminologists, have been pursuing this line of research for quite some time.
- *Individual realities of marginalized peoples:* A key aspect of empirical postmodern research is its focus on uncovering the stories of those who do not usually, if ever, get to speak about their experiences. Through this line of research, we have learned a great deal about street gangs, homeless people, prison inmates, ethnic and racial minorities, women, and the poor. This line of research has helped expand the dominant narrative to be more inclusive and more accurate. For example, the dominant U.S. narrative that everyone who works hard will live a comfortable life has now been shown to be less than accurate, especially when considering the plight of women, some ethnic and minority groups, and others who do not conform to

traditional and accepted images of the hardworking American. Indeed, thanks to the application of postmodernist ideas to empirical research, we can now safely say that *not* everyone who works hard will live a comfortable life.

- *Misunderstanding crime:* Many of the ideas and methodologies derived from postmodernism have uncovered not only that most people do not understand the seriousness or prevalence of crime, but that their attitudes toward crime, offenders, and victims tend to be influenced primarily by their own experiences, their discussions with others, and distorted and inaccurate media portrayals of crime and criminal justice. Unfortunately, it appears that hard data rarely seem to influence people's opinions about crime. The policy implications of this discovery are already being felt in this country as more than 2 million people continue to be incarcerated in our jails and prisons and more and more of our tax dollars go to building prisons, hiring additional police, and funding other types of criminal justice activity.
- *Atomization, social capital, and collective efficacy:* Finally, and perhaps most important, postmodernism and the conditions of postmodernity (or radical modernity, as Giddens argues) suggest that as trust erodes, we will become focused inward, less interested in collective goals, and less willing to work with others to solve problems unless the problems are directly related to our own needs. The implications of this condition are clear for a democratic system of governance: It will cease to function as it was intended, and powerful interest groups (and other elites) will dictate the terms of our democratic system.

Future Directions for Research

As should be clear to the reader by now, postmodernism offers an unlimited number of possibilities for future research. Some of the ideas and methodologies influenced by postmodernism (constructivism in particular) are content analyses, ethnomethodologies (institutional ethnography, ethnographic realism, confessional ethnography, critical ethnography, dramatic ethnography), narrative analyses, historical narratives, life histories, ethnographic case studies, focus groups, and other forms of qualitative research that seek out "thick descriptions" rather than generalizable, statistical knowledge. Future research based on any one of these methodologies is entirely up to the individual researcher. Whether it be philosophical, theoretical, or applied research, the number of topics at the researcher's fingertips is limitless.

That being said, a great deal of research effort continues to focus on the role of the mass media in our society. In

particular, researchers continue to pursue knowledge about the influence of the mass media on policymakers and public policy, their influence on people's beliefs about any given policy, their role in agenda setting, their influence on the behavior of people (e.g., video games and juvenile crime), and other areas in which the mass media are influential. Recently, attention has begun to focus on the new media, or Internet sources of information such as blogs, social network sites, advertising, Internet porn, and other Internet phenomena. Much more needs to be done in this arena as the old theories and understandings regarding the role of the media in society may not hold up when one is attempting to understand the influence of the Internet on society.

Conclusion

In this chapter, the broad topic of postmodernism was summarized and distilled into its essential elements. In discussing the essential literature regarding the assumptions of postmodernism, it is hoped that students will become intrigued and inspired by the ideas put forward by the founders of postmodern thought. The application of postmodern thought to empirical research holds a number of interesting possibilities—possibilities that do not necessarily need to supplant quantitative approaches but rather may supplement these approaches with a more complete accounting of subjects that are difficult to measure (in a quantitative sense) and are rarely the focus of positivist research. Because postmodernism offers students of the social sciences alternatives to positivism, it expands opportunities for substantive, empirical research on a variety of intriguing topics. Conducting research from a postmodern perspective is, in many ways (and as many postmodernists would argue), the route to intellectual freedom. Simply, the only limits placed on the postmodernist researcher are the limits of the researcher's imagination.

The policy implications offered by postmodern thought and research are significant. In particular, the notion that reality is entirely subjective and that there is a dark side to our existence in postmodern bureaucratic systems suggests a number of interesting research topics that have significant implications for public policy, organizational theory, criminology, and other areas of social science research. Important in any research agenda based on postmodernist assumptions is the notion that dominant narratives and discourse blot out myriad politically and socially important experiences, perspectives, and voices of both individuals (e.g., prison inmates, radical political actors, sex offenders) and entire groups of people (e.g., undocumented immigrants, women, ethnic and racial minority groups). Simply, postmodernists draw our attention to concerns that researchers would not otherwise consider or have considered from dominant perspectives and methodologies that intentionally (or unintentionally) obscure the day-to-day reality of marginalized groups.

An engaging example of the application of postmodern assumptions to empirical scholarship is the work of sociologist Sudhir Venkatesh (2008). Venkatesh wrote *Gang Leader for a Day: A Rogue Sociologist Takes to the Streets* after nearly a decade observing gang behavior from inside the gang. Although we cannot generalize from his research, it substantively contributes to our knowledge of urban poverty, the day-to-day activities occurring in a Chicago housing project, and violent street gangs (i.e., marginalized groups and individuals). This is not work that could have been accomplished (or accomplished as thoroughly) through positivist methods.

As is the case with much postmodern scholarship, Dr. Venkatesh stumbled into his line of research. Indeed, much postmodern research is a result of accident, chance, and the ability of postmodern researchers to expand (and implement) their sociological imagination (Mills, 1959). The task of the postmodernist is to move beyond the obvious and measurable and look behind the scenes at those phenomena that cannot be easily measured or quantified. Postmodern thought, then, is an important epistemological, ontological, and methodological step toward developing a more complete understanding of the confusing, diverse, and busy world that we live in today.

References and Further Readings

- Adams, G. B., & Balfour, D. L. (1998). *Unmasking administrative evil*. London: Sage.
- Agamben, G. (1998). *Homo sacer: Sovereign power and bare life* (D. Heller Roazen, Trans.). Stanford, CA: Stanford University Press.
- Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Wadsworth.
- Baudrillard, J. (1981). *Simulacra and simulations* (P. Foss, P. Patton, & P. Beithman, Trans.). New York: Semiotext(e).
- Baudrillard, J. (1996). *The perfect crime* (C. Turner, Trans.). London: Verso.
- Bergquist, W. (1993). *The postmodern organization: Mastering the art of irreversible change*. San Francisco: Jossey Bass.
- Bohm, R. M. (2006). "McJustice": On the McDonaldization of criminal justice. *Justice Quarterly*, 23(1), 127-146.
- Clear, T. (2007). *Imprisoning communities: How mass incarceration makes disadvantaged communities worse*. Oxford, UK: Oxford University Press.
- Crank, J. P. (2003). *Imagining justice*. Cincinnati: Anderson.
- Derrida, J. (1997). *Of grammatology* (G. C. Spivak, Trans.). Baltimore: Johns Hopkins University Press. (Original work published 1967)
- Derrida, J., & Bass, A. (2004). *Position*. (A. Bass, Trans.). New York: Continuum.
- de Zwart, F. (2002). Administrative practice and rational inquiry in postmodern public administration theory. *Administration & Society*, 34(5), 482-498.
- Farmer, J. D. (1997). Leopards in the temple: Bureaucracy and the limits of the in between. *Administration & Society*, 29(5), 507-529.
- Feeley, M. M., & Simon, J. (1992). The new penology: Notes on the emerging strategy of corrections and its implications. *Criminology*, 30(4), 449-474.
- Foucault, M. (1980). *Power/knowledge: Selected interviews & other writings 1972-1977* (C. Gordon, Ed.; C. Gordon, L. Marshall, J. Mepham, & K. Soper, Trans.). New York: Pantheon Books.
- Foucault, M. (1995). *Discipline and punish: The birth of the prison*. New York: Vintage Books.
- Garland, D. (2001). *The culture of control: Crime and social order in contemporary society*. Chicago: University of Chicago Press.
- Giddens, A. (1990). *The consequences of modernity*. Stanford, CA: Stanford University Press.
- Gordon, D. R. (1990). *The justice juggernaut: Fighting street crime, controlling citizens*. New Brunswick, NJ: Rutgers University Press.
- Grahame, P. R. (1998). Ethnography, institutions, and the problematic of the everyday world. *Human Studies*, 21, 347-360.
- Harding, S. (Ed.). (2004). *The feminist standpoint theory reader: Intellectual and political controversies*. New York: Routledge.
- Irwin, J. (1985). *Jail: Managing the underclass in American society*. Berkeley: University of California Press.
- Irwin, J. (2005). *The warehouse prison: Disposal of the new dangerous class*. Los Angeles: Roxbury.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Lawrence, R. (2000). *The politics of force: Media and the construction of police brutality*. Berkeley: University of California Press.
- Lynch, M. (1998). Waste managers? The new penology, crime fighting, and parole age identity. *Law & Society Review*, 32(4), 839-869.
- Lyotard, J. F. (1979). *The postmodern condition*. Manchester, UK: Manchester University Press.
- Miller, L. (2001). Looking for postmodernism in all the wrong places: Implementing a new penology. *British Journal of Criminology*, 41, 168-184.
- Mills, C. W. (1959). *The sociological imagination*. New York: Oxford University Press.
- O'Neill, J. (Ed.). (1996). *Hegel's dialectic of desire and recognition: Texts and commentary*. Albany: SUNY Press.
- Selznick, P. (1992). *The moral commonwealth: Social theory and the promise of community*. Berkeley: University of California Press.
- Simon, J. (1997). Governing through crime. In L. Friedman & G. Fisher (Eds.), *The crime conundrum: Essays on criminal justice* (pp. 171-189). Boulder, CO: Westview.
- Smith, D. E. (Ed.). (2006). *Institutional ethnography as practice*. Lanham, MD: Rowman & Littlefield.
- Surette, R. (2007). *Media, crime, and criminal justice: Images, realities, and policies* (3rd ed.). Belmont, CA: Thomson Wadsworth.
- Venkatesh, S. (2008). *Gang leader for a day: A rogue sociologist takes to the streets*. New York: Penguin Books.
- Weber, M. (1946). Bureaucracy. In H. A. Gerth & C. W. Mills (Eds. & Trans.), *From Max Weber: Essays in sociology* (pp. 196-244). New York: Oxford University Press.
- Welch, M. (2005). *Ironies of imprisonment*. Thousand Oaks, CA: Sage.

3

NEOINSTITUTIONALISM

MARK C. MILLER

Clark University

Neoinstitutionalism, also known as the *new institutionalism*, has been one of the primary methodological approaches in political science in the United States since the late 1980s. This methodology is especially popular among scholars of U.S. politics, although it is growing in influence in the fields of comparative politics and international relations. The new institutionalism combines the interests of traditionalist scholars in studying formal institutional rules and structures with the focus of behavioralist scholars on examining the actions of individual political actors. The new institutionalism thus explores how institutional structures, rules, norms, and cultures constrain the choices and actions of individuals when they are part of a political institution. In other words, “The neo-institutionalist perspective combines the microlevel study of individual behavior with the macrolevel sensitivity to the institutional factors that help shape that behavior” (Miller, 1995, p. 6). The new institutionalism is a very influential postbehavioralist methodology today among political scientists in the United States and abroad.

The Historical Roots of the New Institutionalism

The Traditionalists

From the 1930s through the 1950s, traditionalist scholars dominated political science as a discipline, and especially

political science as practiced in the United States. These scholars were most interested in examining the formal structures and rules that were the foundation of political and governmental institutions such as the executive branch, the legislative branch, and the judiciary. As Rhodes, Binder, and Rockman (2006) explain, “When political science emerged as a separate field, it emphasized the study of formal-legal arrangements as its exclusive subject matter” (p. xii). These studies were often descriptive in nature, using mostly qualitative methods, and they usually did not use broad theories in order to ground their observations in a larger theoretical perspective. Often they were quite normative in their desire to describe how political institutions ought to function, as opposed to the empirical study of how things actually worked in practice. Rhodes et al. thus describe the traditionalist approach in this way:

The older studies of institutions were rooted in law and legal institutions, focusing not only on how “the rules” channeled behavior, but also on how and why the rules came into being in the first place, and, above all, whether or not the rules worked on behalf of the common good. (p. xii)

The Behavioral Revolution

Beginning in the 1960s, political scientists began to move away from focusing on political institutions and instead almost exclusively studied the actions of individual

political actors. This so-called behavioral or behaviorist revolution thus focused on making the study of politics more scientific, and quantitative methods came to predominate in political science. The behaviorist revolution was especially critical for students of U.S. politics. Since good quantitative studies demanded large sample sizes, the more qualitative studies of single institutions and institutional rules waned in part because of their small sample sizes. For example, instead of studying the structures and rules of the courts, behaviorist political scientists studied specific decisions of individual judges. Or instead of studying the role of Congress in the broader system of government, behaviorists instead studied the choices made by individual members of Congress or by the voters in congressional elections. The hope was that political scientists would develop broad theoretical approaches that would be validated by quantitative empirical methods, thus moving political science away from the disciplines of history, law, and philosophy and instead bringing it closer to the scientific approaches of economics, sociology, and psychology. In some ways, the behaviorist revolution privileged those who wanted political science to be more like a hard science over those who favored so-called softer approaches to the study of politics. Behaviorists stressed rigorous empirical analysis of the behavior of individual political actors.

By the mid-1980s, many political scientists began to question whether the discipline should continue to ignore the traditionalist interest in political institutions but without abandoning what we had learned from the behaviorist approach in examining the choices of individuals. There was a worry that behavioralism could bring us only so far and that perhaps we had learned all we could from that approach. Therefore a so-called postbehaviorist movement arose within political science, designed in part to bring the study of institutions back into the discipline. It seemed natural to many scholars that political scientists would study political institutions again. As Rhodes et al. (2006) argue, "The study of political institutions is central to the identity of the discipline of political science" (p. xii).

The New Institutionalism Emerges

The new institutionalist approach has its roots in the early to mid-1980s. Often considered two of the leading founders of the new institutionalism, James G. March and Johan P. Olsen published a very influential piece in the *American Political Science Review* in 1984 titled "The New Institutionalism: Organizational Factors in Political Life," followed by a book published in 1989 titled *Rediscovering Institutions*. They continued to argue for further institutional analysis in their 1995 book, *Democratic Governance*. In all these now classic pieces, these scholars argued the then radical position that political scientists needed to rediscover institutional analysis in order to understand better the behavior of individual political actors within political institutions. In other words, according to

these authors, studying individual political behavior without examining institutional constraints on that behavior was giving scholars a skewed understanding of political reality. Thus, March and Olsen argued that studying institutions again would allow political scientists to discover more of the complexities of politics. As these authors have concluded in a more recent work, "Institutions empower and constrain actors differently and make them more or less capable of acting according to prescriptive rules and appropriateness" (March & Olsen, 2006, p. 3).

By the mid- to late 1990s, the new institutionalism came to be one of the dominant approaches in political science, especially among those who studied U.S. politics. Today it is also used by many scholars in comparative politics and in international relations as well. Referring to what was happening among scholars in the late 1980s and throughout the 1990s, Barry Weingast (2002) has stated, "Political science witnessed a revolution in the study of institutions" (p. 660). Institutional analysis was being used to study everything from the legislative process to effects of social movements and to the politics of the judiciary. The new institutionalist approach has become so influential that Pierson and Skocpol (2002) could claim that "we are all institutionalists now" (p. 706).

A Return to the Law and Legal Analysis

Since traditionalist political science had its roots in the approaches of law and philosophy, perhaps it is not surprising that many scholars in the field of judicial politics have strongly embraced the new institutionalism. While the behaviorists, using mostly quantitative methods, often studied the decision-making processes of individual judges, some political scientists continued to argue that legal rules, structures, and doctrines still mattered. After a long period in which the behaviorist approach dominated the study of judicial politics, Rogers Smith (1988) was one of the first to argue that public law scholars should use a more institutional approach. As Smith (2008) has argued in a more recent work, "It took no great insight to realize that these emphases on the importance of rules in bounding action and constituting actors, while simultaneously enabling rule-interpreters to make choices that shaped outcomes, might aid scholars of law and courts" (p. 48). In the late 1980s, many judicial politics scholars had already started using new institutionalist methods. By the late 1990s, the new institutionalism approach in judicial politics had become so popular that Howard Gillman and Cornell Clayton in 1999 edited two highly influential volumes of essays on new institutionalist studies of decision making on the U.S. Supreme Court. About this same time, new institutionalist approaches were also becoming popular for scholars of U.S. legislatures (see, e.g., Miller, 1995), of the U.S. presidency (see, e.g., Skowronek, 1993), of various aspects of comparative politics (see, e.g., Hall, 1986), and of international relations (see, e.g., Keck & Sikkink, 1998).

The New Institutionalism in Comparative Politics

Just as the new institutionalism was becoming important in the study of U.S. politics and law, it was also becoming a feature of much of the work in the study of comparative political institutions. In the late 1960s, Roy C. Macridis (1968) called on comparative politics scholars to abandon the traditionalist approach in that field, which he termed “parochial, monographic, descriptive, bound to the West, . . . excessively formalistic and legalistic, and insensitive to theory-building and theory-testing” (p. 79). His solution was an early call for postbehavioralist analysis of “governmental institutions and political elites, their role, their levels of performance and nonperformance” (p. 89). By this, he meant comparative scholars should focus on parliaments, the executive, the judiciary, the civil service, and other aspects of governmental institutions. Macridis’s efforts were in part a precursor of what happened in international relations in the early 1980s, when various scholars called for “bringing the state back in” to international theory building (see, e.g., Evans, Rueschemeyer, & Skocpol, 1985). Some of the new institutionalist analysis in comparative politics chose a “thin” approach, which looked at the role of institutions in various public policy issues in various settings. Others took a “thick” approach, which added governmental institutions to the previous analysis of other social structures, such as social movements, political coalitions, and ideological constraints within a society (see, e.g., Kohli et al., 1995). In other words, should the study of governmental institutions be separate from cultural studies or be used in combination with cultural approaches (see, e.g., Lecours, 2000)?

Especially since the 1990s, new institutional analysis has been used in a variety of comparative politics arenas. One important work in this area was an edited volume titled *Do Institutions Matter? Government Capabilities in the United States and Abroad* (Weaver & Rockman, 1993). This book covered a wide range of issues, from specific public policies such as energy policy to foreign affairs and to separation-of-powers issues. Other scholars used new institutional approaches to study such public policies as health, welfare, and industrial development (see, e.g., Steinmo, Thelen, & Longstreth, 1992). Some scholars studied the role of governmental institutions in the economy more generally (see, e.g., Hall, 1986). Others studied the role of international organizations such as the European Union (see, e.g., Pierson, 1996). Other scholars studied comparative legislative institutions; comparative leadership, mostly in the executive branch of government; comparative courts; or comparative electoral systems. Other new institutionalist scholars studied topics that many do not immediately associate with political institutions, such as comparative revolutions or political conflict in general (see, e.g., Peters, Pierre, & King, 2005). Another set of scholars began to examine informal institutions, such as political norms and rules that are “created, communicated, and enforced outside of officially sanctioned

channels” (Helmke & Levitsky, 2004, p. 725). These informal rules can have serious political ramifications in many societies. Other scholars are using new institutionalist analysis to examine concepts such as state formation.

Is There One Definition of a Political Institution?

While many political scientists now use a new institutionalist approach, there has been little agreement among these scholars on a single precise definition of what constitutes a political institution. As Rhodes et al. (2006) note, “Despite the incredible growth in institutional studies in recent decades, we lack a singular definition of an institution on which students of politics can find wide agreement” (p. xiii). All seem to agree that a political institution is an entity of its own and that an institution is more than just the sum of the policy preferences of the individuals who comprise it. For example, Carey (2000) argues that political institutions “establish guidelines for deliberation, the aggregation of preferences into collective decisions, and the implementation of those decisions” (p. 735). And all seem to agree that an institution is a concept more than a place or a thing. Thus Rawls (1971) conceptualized political institutions as “an abstract object” realized in “thought and conduct” (p. 11). March and Olsen (2006) have provided this rather complicated definition of what they study:

An institution is a relatively enduring collection of rules and organized practices, embedded in structures of meaning and resources that are relatively invariant in the face of turnover of individuals and relatively resilient to the idiosyncratic preferences and expectations of individuals and changing external circumstances. (p. 3)

Thus, according to these scholars, “Institutions give order to social relations, reduce flexibility and variability in behavior, and restrict the possibilities of a one-sided pursuit of self-interest or drives” (March & Olsen, 2006, p. 7).

Although political scientists cannot seem to agree on a definition of an institution, it has been easier for many scholars to state what institutions are not. Thus Brigham (1987) has noted that “institutions are not simply robes and marble, nor are they contained in codes or documents” (p. 21). And Howard Gillman (1999), in his institutional approach to the study of decision making on the U.S. Supreme Court, titles one of his chapters “The Court as an Idea, not a Building (or a Game)” (p. 65).

Multiple Levels of Analysis

One of the advantages of new institutionalist approaches is the fact that the new institutionalism allowed scholars to use multiple levels of analysis simultaneously. The

traditionalists focused mainly on macrolevel analysis, using institutions as the unit of analysis. Thus the traditionalists studied institutions qua institutions. The behavioralists, on the other hand, focused almost exclusively on microlevel analysis, because understanding the actions of individual political actors was their primary concern. The new institutionalists, however, said that scholars could and should examine multiple levels of analysis, often within the same research project. Thus, for example, for congressional scholars, the unit of analysis could be the individual member of Congress, or a congressional committee, or the party caucuses, or congressional leadership, or a chamber of Congress, or the Congress as a whole (see, e.g., Miller, 1995). This ability to examine multiple levels of analysis simultaneously also enabled new institutionalist scholars to explore the interactions between and among political institutions, giving rise to many studies concerning, for example, the relationship between U.S. legislatures and the courts (see, e.g., Miller & Barnes, 2004), between Congress and the presidency (see, e.g., Fisher, 1981), between the bureaucracy and the Congress (see, e.g., Ripley & Franklin, 1980), and so forth. As Scheingold (2008) concludes, one of the essential messages of these studies is “the fluid and reciprocal character of institutional interaction” (p. 742).

Three Streams of New Institutionalism

One of the reasons that there is no single agreed-on definition of a political institution is that the new institutionalist approach is really an umbrella term for a wide variety of complementary but clearly different methodologies (see, e.g., Thelen, 1999). In their now classic discussion of the various strains of new institutionalist analysis within political science, Hall and Taylor (1996) argue that there are at least three branches of new institutionalism: *rational choice institutionalism*, *sociological institutionalism*, and *historical institutionalism*. Rational choice institutionalism has its roots in economics and formal modeling analysis. Sociological institutionalism has its roots in sociology, anthropology, and cultural studies. And historical institutionalism has its roots in the disciplines of history and law. Scholars such as Hay (2006) and Ansell (2006) want to further subdivide these three branches by inventing new labels such as *constructionist institutionalism* and *net work institutionalism*. Schmidt (2008) uses the term *discursive institutionalism*. But most scholars focus more on the three main streams of new institutionalist analysis.

Rational Choice Institutionalism

Rational choice institutionalism is one of the major approaches within the new institutionalist umbrella, with its roots in economic analysis and formal theory. A popular approach within this stream is the use of game theory

to explain political decision making (see, e.g., Shepsle, 2006). Thus, according to Rhodes et al. (2006), “Rational-choice institutionalists think of institutions as a system of rules and incentives” (p. xiii). These rules are often contested so that one set of political actors can gain an advantage over a different group. Rhodes et al. continue:

Institutions in this sense provide arenas for conflict, and efforts to alter them stimulate conflict inasmuch as they change the rules of the game in such a way as to alter the allocation of advantages and disadvantages. From this vantage point rules are never neutral, but are instead part of a struggle between challengers and holders of power. (p. xiv)

Thus rational choice scholars often focus on a single institution in a specific time frame, although more and more are looking at institutions across time.

This conceptualization of institutions as arenas for game playing underlies much of the formal modeling done by rational choice scholars. Thus the rational choice scholars are simultaneously theoretical and empirical. The formal models produced by rational choice scholars attempt to simplify the political world in order to explain its essential features. According to Weingast (2002), these formal models allow rational choice scholars to study “how institutions constrain the sequence of interaction among the actors, the choices available to particular actors, the structure of information and hence beliefs of the actors, and the payoffs to individuals and groups” (p. 661). These models help answer questions such as why political institutions are needed in the first place, why they take on particular forms, and why they survive over time. This approach puts a great deal of stress on concepts such as efficiency and rationality of decision making. As Shepsle (2006) explains this approach, “The research program of rational choice institutionalism is founded on abstraction, simplification, analytical rigor, and an insistence on clean lines of analysis from basic axioms to analytical propositions to empirical implications” (p. 32).

A subset of rational choice institutionalism has come to be known as the *strategic approach*. This approach is especially popular among judicial scholars. Accepting the new institutionalist idea that institutions constrain the behavior of individual political actors, these scholars argue that individuals within political institutions often act strategically, thus anticipating the reactions of their colleagues to their decisions, as well as anticipating the reactions of other institutions to their decisions. Instead of pushing for their ultimate policy preferences, these actors temper their actions by anticipating the reactions of others. Thus their votes and actions are constrained by their understanding of institutional factors and by the anticipated reactions of other actors and other institutions to their decisions. As Hall and Taylor (1996) conclude, “One of the great contributions of rational choice institutionalism has been to emphasize the role of strategic interaction in the determination of political outcomes” (p. 945).

Sociological Institutionalism

A second branch of new institutionalist analysis is the sociological approach. It has its roots in organizational theory, anthropology, and cultural studies. This stream stresses the idea of institutional cultures. These scholars see institutional rules, norms, and structures not as inherently rational or dictated by efficiency concerns, but instead as culturally constructed. As Hall and Taylor (1996) explain, sociological institutionalists argue that “even the most seemingly bureaucratic of practices have to be explained in cultural terms” (p. 947). These scholars tend to define institutions more broadly than do scholars in the other two streams. These scholars tend to look at the role of myth and ceremony in creating institutional cultures, as well as the role of symbol systems, cognitive scripts, and moral templates. At times these scholars take on a normative approach to the study of political institutions, and they tend to blur the line between institutions and culture. Their work often focuses on questions of the social and cultural legitimacy of the organization and its participants. The pioneering work in this field was done by sociologists at Stanford University (see, e.g., Powell & DiMaggio, 1991). This branch of new institutionalist analysis is probably the least influential among contemporary political scientists, although many use the concepts of institutional culture and institutional will in their research projects.

Historical Institutionalism

The third branch of new institutionalist analysis, historical institutionalism, has received a great deal of attention among political scientists, especially those who use more qualitative methodologies in studying U.S. politics. But it is also becoming more popular in comparative politics and in studies of international relations. In some ways, historical institutionalism is the hardest of the three branches to define because it includes so many different scholars and so many different methodological approaches. This branch includes an eclectic group of scholars with a wide variety of research agendas (see, e.g., Lecours, 2000; Thelen, 1999). Despite their differences, there are some common notions in this line of research. As Pierson and Skocpol (2002) argue, within this group of scholars, “Everyone seems to realize that theoretical eclecticism, multiple analytic techniques, and a broad comparative and historical purview work best” (p. 698).

Generally, historical institutionalists define political institutions broadly and are interested in changes in institutions over time. For example, Hall and Taylor (1996) argue that historical institutionalists define institutions as “the formal or informal procedures, routines, norms, and conventions embedded in the organizational structure of the polity or political economy” (p. 938). Rhodes et al. (2006) note that historical institutionalists “see institutions as continuities” (p. xv). Thus, according to these scholars,

political institutions may be constitutional in nature, procedural, or programmatic. In comparative politics, the new institutionalist scholars compare policies not only across time but also across countries (Peters et al., 2005).

Historical institutionalists think a lot about *decision trees* and *path dependence*, terms of art meaning the effects that one decision has to limit the available future choices for any political actor or institution (see, e.g., Peters et al., 2005). While historical institutionalists clearly focus on institutional analysis, these scholars “rarely insist that institutions are the only causal force in politics” (Hall & Taylor, 1996, p. 942). Thus, for example, historical institutionalists acknowledge the importance of ideas in creating political change, as well the role of economic or cultural variables in political decision making. They also attempt to identify critical junctures in the political process, as well as the goals of political actors. Some of these studies are inherently comparative, whereas others focus on a single political institution. As Sanders (2006) argues, for historical institutionalists, “What is mainly of interest is the construction, maintenance, and adaptation of institutions” (p. 42).

While it is difficult to pinpoint a precise definition of the historical institutionalist approach, there are some common traits in this line of scholarship. Pierson and Skocpol (2002) have listed three important characteristics of this approach. First, they note that historical institutionalists “address big, substantive questions that are inherently of interest to broad publics as well as to fellow scholars” (p. 695). Thus the research puzzles of these scholars are often rooted in real-world political problems. Second, these historical institutionalists “take time seriously” (p. 695), and thus they can trace changes in politics and institutions over history. Extending the time frame that political scientists consider gives them more empirical data to analyze and allows an examination of events that occur rather rarely, such as democratization and revolutions. It also allows scholars to explore lengthy, large-scale, but slow-moving social processes in which change often occurs only incrementally. Some of the focus of these scholars is thus on institutional development issues (see, e.g., Orren & Skowronek, 2002). Third, historical institutional scholars pay clear attention to the contexts and configurations that allow them to “hypothesize about the combined effects of institutions and processes” (Pierson & Skocpol, 2002, p. 696). In other words, historical institutionalists usually do not examine only one institution or process at a set point in time, but rather they tend to look at politics as a very complex set of processes and institutions that vary over time and that interact in interesting and unexpected ways. As Pierson and Skocpol (2002) conclude, “The focus is on explaining variations in important or surprising patterns, events, or arrangements—rather than on accounting for human behavior without regard to context or modeling very general processes presumed to apply at all times and places” (pp. 696–697).

Conclusion

The new institutionalist approach in political science remains a very influential methodological tool. Many scholars consider themselves to be new institutionalists, and there is no one precise definition of what constitutes the new institutionalism. There are at least three streams within this broad approach, and some scholars borrow ideas and concepts from multiple streams in order to answer their research questions. The one thing that all these scholars have in common is that they take institutions seriously because they feel that political institutional factors constrain the choices of individual political actors. In other words, the new institutionalism in political science is a postbehavioralist methodological approach that stresses the importance of political institutions in furthering our understanding of politics and government.

References and Further Readings

- Ansell, C. (2006). Network institutionalism. In R. A. W. Rhodes, S. A. Binder, & B. A. Rockman (Eds.), *The Oxford handbook of political institutions* (pp. 75–89). Oxford, UK: Oxford University Press.
- Brigham, J. (1987). *The cult of the court*. Philadelphia: Temple University Press.
- Carey, J. (2000). Parchment, equilibria, and institutions. *Comparative Political Studies*, 33, 735–761.
- Clayton, C., & Gillman, H. (Eds.). (1999). *Supreme Court decision making: New institutionalist approaches*. Chicago: University of Chicago Press.
- Evans, P. B., Rueschemeyer, D., & Skocpol, T. (Eds.). (1985). *Bringing the state back in*. New York: Cambridge University Press.
- Fisher, L. (1981). *The politics of shared power: Congress and the executive*. Washington, DC: CQ Press.
- Gillman, H. (1999). The Court as an idea, not a building (or a game): Interpretive institutionalism and the analysis of Supreme Court decision making. In C. Clayton & H. Gillman (Eds.), *Supreme Court decision making: New institutionalist approaches* (pp. 65–90). Chicago: University of Chicago Press.
- Gillman, H., & Clayton, C. (Eds.). (1999). *The Supreme Court in American politics: New institutionalist interpretations*. Lawrence: University Press of Kansas.
- Hall, P. A. (1986). *Governing the economy: The politics of state intervention in Britain and France*. New York: Oxford University Press.
- Hall, P. A., & Taylor, R. C. R. (1996). Political science and the three new institutionalisms. *Political Studies*, 44, 936–957.
- Hay, C. (2006). Constructionist institutionalism. In R. A. W. Rhodes, S. A. Binder, & B. A. Rockman (Eds.), *The Oxford handbook of political institutions* (pp. 56–74). Oxford, UK: Oxford University Press.
- Helmke, G., & Levitsky, S. (2004). Informal institutions and comparative politics: A research agenda. *Perspectives on Politics*, 2, 725–740.
- Keck, M., & Sikkink, K. (1998). *Activists beyond borders: Advocacy networks in international politics*. Ithaca, NY: Cornell University Press.
- Kohli, A., Evans, P., Katzenstein, P. J., Praeworski, A., Rudolph, S. H., Scott, J. C., et al. (1995). The role of theory in comparative politics: A symposium. *World Politics*, 48, 1–49.
- Lecours, A. (2000). Theorizing cultural identities: Historical institutionalism as a challenge to the culturalists. *Canadian Journal of Political Science*, 33, 499–522.
- Macridis, R. C. (1968). Comparative politics and the study of government: The search for focus. *Comparative Politics*, 1, 79–90.
- March, J. G., & Olsen, J. P. (1984). The new institutionalism: Organizational factors in political life. *American Political Science Review*, 78, 734–749.
- March, J. G., & Olsen, J. P. (1989). *Rediscovering institutions*. New York: Free Press.
- March, J. G., & Olsen, J. P. (1995). *Democratic governance*. New York: Free Press.
- March, J. G., & Olsen, J. P. (2006). Elaborating the “New Institutionalism.” In R. A. W. Rhodes, S. A. Binder, & B. A. Rockman (Eds.), *The Oxford handbook of political institutions* (pp. 3–22). Oxford, UK: Oxford University Press.
- Miller, M. C. (1995). *The high priests of American politics: The role of lawyers in American political institutions*. Knoxville: University of Tennessee Press.
- Miller, M. C., & Barnes, J. (2004). *Making policy, making law: An interbranch perspective*. Washington, DC: Georgetown University Press.
- Orren, K., & Skowronek, S. (2002). The study of American political development. In I. Katznelson & H. V. Miller (Eds.), *Political science: State of the discipline* (pp. 722–754). New York: W. W. Norton.
- Peters, B. G., Pierre, J., & King, D. S. (2005). The politics of path dependency: Political conflict in historical institutionalism. *Journal of Politics*, 67, 1275–1300.
- Pierson, P. (1996). The path to European integration: A historical institutionalist analysis. *Comparative Political Studies*, 29, 123–163.
- Pierson, P., & Skocpol, T. (2002). Historical institutionalism in contemporary political science. In I. Katznelson & H. V. Miller (Eds.), *Political science: State of the discipline* (pp. 693–721). New York: W. W. Norton.
- Powell, W. W., & DiMaggio, P. J. (Eds.). (1991). *The new institutionalism in organizational analysis*. Chicago: University of Chicago Press.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Rhodes, R. A. W., Binder, S. A., & Rockman, B. A. (Eds.). (2006). *The Oxford handbook of political institutions*. Oxford, UK: Oxford University Press.
- Ripley, R., & Franklin, G. (1980). *Congress, the bureaucracy, and public policy*. Homewood, IL: Dorsey Press.
- Sanders, E. (2006). Historical institutionalism. In R. A. W. Rhodes, S. A. Binder, & B. A. Rockman (Eds.), *The Oxford handbook of political institutions* (pp. 39–55). Oxford, UK: Oxford University Press.
- Scheingold, S. A. (2008). The path of the law in political science: De centering legality from olden times to the day before yesterday. In K. E. Whittington, R. D. Kelemen, & G. A. Caldeira (Eds.), *The Oxford handbook of law and politics* (pp. 737–751). Oxford, UK: Oxford University Press.
- Schmidt, V. A. (2008). Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science*, 11, 303–326.
- Shepsle, K. A. (2006). Rational choice institutionalism. In R. A. W. Rhodes, S. A. Binder, & B. A. Rockman (Eds.),

28 • GENERAL APPROACHES IN POLITICAL SCIENCE

- The Oxford handbook of political institutions* (pp. 23–38). Oxford, UK: Oxford University Press.
- Skowronek, S. (1993). *The politics presidents make: Leadership from John Adams to Bill Clinton*. Cambridge, MA: Belknap Press.
- Smith, R. M. (1988). Political jurisprudence, the “new intuitionism,” and the future of public law. *American Political Science Review*, 82, 89–108.
- Smith, R. M. (2008). Historical institutionalism and the study of law. In K. E. Whittington, R. D. Kelemen & G. A. Caldeira (Eds.), *The Oxford handbook of law and politics* (pp. 46–62). Oxford, UK: Oxford University Press.
- Steinmo, S., Thelen, K., & Longstreth, F. (Eds.). (1992). *Structuring politics: Historical institutionalism in comparative analysis*. Cambridge, UK: Cambridge University Press.
- Thelen, K. (1999). Historical institutionalism in comparative politics. *Annual Review of Political Science*, 2, 369–404.
- Weaver, R. K., & Rockman, B. A. (Eds.). (1993). *Do institutions matter? Government capabilities in the United States and abroad*. Washington, DC: Brookings.
- Weingast, B. R. (2002). Rational choice institutionalism. In I. Katznelson & H. V. Milner (Eds.), *Political science: State of the discipline* (pp. 660–692). New York: W. W. Norton.

4

SYSTEMISM

HYUNSUN CHOI

University of North Florida

Systemism has emerged as an important worldview and methodological approach in social science. This approach is generally against reductionism, and it sees everything either as a system or as part of a system. This view is different from individualism or holism. While individualism emphasizes individuals in society, holism focuses on structure. Systemism can be seen as an alternative way to make sense of a complex world.

This chapter explores the historical and theoretical development of the systemism approach in social science by addressing its applications and policy implications. Systemism contributes to methodological issues such as systems analysis, modeling, case study, and survey research, and it may have significant policy implications in the fields of environmental politics, administrative decision making, and urban politics and development.

The Evolution of Systemism Theory

In social science studies, there are generally three different broad perspectives to understanding behavior: individualism, holism, and systemism. Systemism can be seen as being situated between individualism and holism.

Individualism

Individualism emphasizes the important role of an individual in society. It claims that society exists for the benefit

of the individual, and the individual must not be constrained by government interventions or made subordinate to collective interests. Ayn Rand, who was a philosopher of the early 20th century, wrote that humans are ultimate ends in themselves, not means to the ends of others. The pursuit of one's own self-interest and happiness is the highest moral purpose of life. Thomas Hobbes, John Locke, Adam Smith, and Max Weber used similar ideas to describe the world in the philosophical tradition of individualism. Social systems are collections of individuals.

In sum, individualism focuses on (a) protecting natural freedoms and rights, (b) pursuing the development of society for the benefit of the individual, and (c) supporting capitalism with minimum government. However, it fails to recognize the causes of social problems and the importance of government intervention.

Holism

While individualism is an important way of thinking about individual choices and behaviors in everyday life, holism emerged to address some of the limitations of individualism. Individualism argues that individuals maintain a primary influence over society. It may embody a degree of validity in explaining society, but it fails to credit the influence that the social environment maintains over our thoughts, beliefs, and opinions. Thus, holism asserts that society must be analyzed as a whole system rather than simply in terms of its individual components.

Holism emphasizes that the primary focus of society should be on the collective unity or greater good of the whole. It may support big government and government intervention for the greatest common good. Thus, holism focuses more on political institutions, and its applications may tend toward totalitarianism. The works of Georg Wilhelm Friedrich Hegel, Auguste Comte, Karl Marx, and Émile Durkheim support holism with an emphasis on social or class structure (Bunge, 2000).

Systemism

Systemism tries to synthesize individualism and holism. It seeks a common ground to view the world from a systems perspective, including micro- and macrolevels. Both individualism and holism have some pitfalls. First, individualists fail to realize the existence of systemic social problems such as poverty, discrimination, and underdevelopment. Second, holism may not see individual actions as the source of social change. Systemism is a moderate approach between these two extremes, and it extracts elements from both approaches. The concept of the system is a key aspect of this worldview. Systems are everywhere, and they have diverse sizes and forms over time and space. Political processes and administrative units are systems, and society, community, and neighborhood are also systems. Systemism is more realistic and therefore more useful than individualism or holism.

According to systemism, everything is either a system or a component of a system. Mario Bunge (2000) described the following postulates of systemism:

1. Everything, whether concrete or abstract, is a system or an actual or potential component of a system.
2. Systems have systemic (emergent) features.
3. All problems should be approached in a system rather than in a sectoral fashion.
4. All ideas should be put together in systems (theories).
5. The testing of anything, whether idea or artifact, assumes the validity of other items, which are taken as benchmarks, at least for the time being.

These rules may imply that nothing should be evaluated as an end itself, but rather as a component of something larger or more meaningful. Systemism analyzes individuals and structures in society dynamically, whereas individualism and holism fail to recognize that there is a middle ground between individual and collective agency and to recognize the interrelatedness of both agency and structure in studying micro-macro questions.

Applications

In the real world, system is everywhere. With the broad definition of system, systemism offers practical applications in diverse fields. Those applications can reduce tensions

between centralism and laissez-faire in politics and government interventions.

Systemism in Social Science

We may think of systemism as a spider and a web. When a spider weaves a web, the spider spins and maintains the web and so creates a structure to help itself survive and prosper. The web shows how agency and structure are intertwined and interdependent. In this way, the differences between systemism and individualism or holism can be understood. Systemism blends two incomplete explanations of the relationships of humans with each other and their environment (Denis, 2003). While people interact with different entities, they deal with large amounts of information in the context of every other interaction. Just like the example of spider and web, people are having diverse interactions based on previous perceptions and interactions. This is an illustration of the interdependent nature of structure and agency that is identified by and explored through systemism.

Adam Smith's *invisible hand theory* is another good example of systemism. Smith's thesis is dependent on an individual's interpretation of collective agency, and therefore the invisible hand is a feature of both human agency and structure. The interrelationship of structure and agency thus occurs within the context of a system. The invisible hand theory aggregates individual agency into a macrolevel benefit. Poverty, for instance, is not necessarily solely the result of individual action or structural malfunction. It may be a combination of different factors from individual bad behavior, coupled with the features of a social structure. While any one person may not experience all poverty-related factors, including contributing factors such as access to quality housing, education, or health care; individual behaviors such as wasteful spending; or unfortunate circumstances such as death, disability, sickness, or chronic unemployment, all those issues occur within the system and have unanticipated consequences.

Urban Politics

Urban politics is an interesting example of the systemism approach. Many scholars have used the systems approach as a framework for the analysis of city politics and policy. David Easton (1965) was one of the first researchers to develop an influential theoretical framework that facilitated an understanding of politics, including urban politics. He applied the open system models of natural science to political science. In this view, political systems are open to influences from their environments. The environment of a political system has inputs and changes that can shape the political process. Political and administrative decision makers need to respond to the environment, and the political system transforms inputs into political and policy outputs.

The political process decides allocations of resources. The result of the decision-makers' processing of inputs yields outputs. The impact of these concrete decision outputs of political actors produces a set of outcomes that send information back into the environment. The process is best viewed as a continuous stream of feedback to the environment that may result in the alteration and creation of inputs that keep cycling into the system (Pelissero, 2003). The systems approach has proved to be a useful way of thinking about the interrelationships of actors, institutions, and the environment of cities.

Case Studies

In studies of urban politics, researchers can focus on one city and employ structured methods of data collection and interpretation to develop and test theories of urban politics. Intensive or ethnographic observations have been used in some classic case studies of cities. For example, the early studies of community power were case studies in which researchers often lived in a community, interviewed residents and experts, analyzed data on local government, and observed the play of local politics. The first-generation urban studies tend to focus on single cities.

The case study method has developed into a comparative method, in which multiple cities are studied or independent case studies are launched in several cities, employing similar methods and research questions. One of the early examples of comparative case studies is Williams and Adrian's (1963) classic study, *Four Cities*. Williams and Adrian studied community power and political process in four cities over time. Another example is the Bay Area research projects that examined multiple cities in the San Francisco area. This study reveals the importance of institutions to the political process. The comparative case study approach has been used to advance theories on city politics, including citizen participation in policy making, political parties, urban regimes, racial and ethnic politics, schools, and moral controversies in cities.

Survey Research

Another approach in systemism is survey research for comparative cross-sectional studies of cities. Researchers began collecting data on a large number of cities and used statistical techniques to assess the determinants of politics or public policies in those cities. In more recent years, the development of new statistical methods and the greater availability of data sets on cities have led to more analysis conducted over time. Computers, advanced statistical software programs, and web-based data sets have increased the opportunities for performing comparative time-series analysis of urban political and policy changes. An example is David R. Morgan and John P. Pelissero's (1980) study of the impact of municipal reform; research on mayors; bureaucratic policy changes; and studies of

city councils, including structural evolution and electoral campaign changes.

Systems Analysis

Systems analysis is a mosaic methodology in decision making. It integrates elements from a variety of disciplines—engineering, sociology, biology, philosophy, psychology, economics, and computer science. Broadly, system analysis makes us look at problems as assemblies of interdependent components. For example, in the context of sewage treatment, if organic sewage is being dumped into a river or lake, it generates an inordinate demand for oxygen. However, oxygen is also needed for bacterial decay, which uses the oxygen-converting organic matter to break down inorganic products. Consequently, dumping tends to deplete the oxygen supply of surface waters. By killing off the bacteria of decay, dumping brings a halt to the aquatic cycle of self-purification. Environmental or sanitation engineers may simply try to domesticate the decay bacteria in a treatment plant, artificially supplying them with sufficient oxygen to accommodate the entering organic materials. Inorganic residues are released, and because they have no oxygen demands, the engineers think the problem is solved.

However, this treatment does not consider the river or lake as a system in nature. The treated sewage becomes rich in inorganic residues of decay—carbon dioxide, nitrates, and phosphates—that support the growth of algae. This chain reaction brings more environmental problems into the natural system. Thus, it is important to take a systemism approach when dealing with these kinds of management and environment issues. There are four basic steps in system analysis: problem formulation, modeling, analysis and optimization, and implementation.

Problem Formulation

This first step may be the most difficult one in analyzing a system where we live and work. This step includes a detailed description of the task and identification of the important variables and their relationships. For instance, in the case of an urban transportation system, one using the systemism approach begins by deciding whether the prime objective is better service, lower cost, less pollution, or something else. Then it is necessary to decide what data are necessary: travel times with different transportation modes; passenger miles by gender, age, race, and income; passenger miles by time and place; and so forth. Last, one must identify key decision makers in the urban area and their motivations.

Modeling

A model is a simpler representation of the real-world problem, designed to help researchers. Models can be

physical reconstructions of the real thing, with constants and variables. The modeler's task is probably more artistic than rigorous, more creative than systematic. Model development needs to find a balance between including all relevant aspects of reality and keeping the model simple enough so that it is in line with current theoretical discussions, computation time, and data availability. Ultimately, the test of a model's quality is how effective it is in helping to solve the original problem. In a model, all components are interconnected. Sometimes, it is necessary to develop a quantified model—a set of mathematical relationships.

Analysis and Optimization

If a model finds the best strategy for resolving a given problem, simulation and sensitivity analysis are important tools for determining the best course of action. These techniques are closely related to the development of computer technology. Simulation allows replicating the actual operation and events of any program in an organization. For example, when one studies patrol-deployment strategy in a police department, police administrators can find simulation models valuable for the following purposes: (a) They facilitate detailed investigation of operations throughout the city, (b) they provide a consistent framework for estimating the value of new technologies, (c) they serve as training tools to increase awareness of the system interaction and consequences resulting from everyday policy decisions, and (d) they suggest new criteria for monitoring and evaluating actual operating systems. When one simulates the dispatch and patrol operations of most urban police departments, incidents are generated throughout the city and distributed randomly in time and space according to observed statistical patterns. Simulation provides a tool to assist in answering a wide range of allocation questions.

Sensitivity analysis helps modelers find the best strategy for solving the original problem. Sensitivity analysis consists of making very small changes in a model to show the extent to which results may be altered because of change in one or a few factors. For example, in real practice, sensitivity analysis can use a small change, such as an 8% decrease in judges or prosecutors or a 3% decrease in police.

Implementation

The last step refers to the procedure by which the results from the model are translated into a set of actions in the real world. However, these four steps seldom occur in perfect sequence, and the systems approach is highly interactive. For instance, the sequence might work out in the following way: formulating the problem, selecting objectives, designing alternatives, collecting data, building models, weighing cost against effectiveness, testing for sensitivity, questioning assumptions, reexamining objectives, looking at new alternatives, reformulating the

problem, selecting different or modified objectives, and so forth.

Policy Implications and Directions in the 21st Century

How do we define the appropriate role of government? Systemism may have a new way of thinking of government's role in the 21st century. For centuries, people have argued whether the proper focus of society should be on the community or the individual. Individualism and holism may be seen as the two extreme views in the real policy world. In the United States, liberal and conservative views have battled over the role of government. Generally, conservatives argue, based on the perspective of individualism, that we must preserve morals and increase personal responsibility instead of the social welfare system. Liberals may argue that government needs to involve itself more in social problems and regulate big business to change government systems. Both views may have good arguments for their points. However, these two views have hardly any common ground. Systemism may offer an alternative way for politics and policy. It can offer a framework to connect the concerns of individualism and holism in order for us to reach our future goals.

What are the policy implications of this approach? One interesting issue is environmental policy and climate change. In national and international politics, climate change is an inevitable issue for the 21st century. Already in the United States, liberals and conservatives, Democrats and Republicans, are battling over this global environmental issue. The result has been gridlock on the issue for more than 10 years now. Can systemism offer some common ground for understanding climate change?

In the climate change debate, some argue that human action is the cause. Others postulate that while humans may have some responsibility for climate change, this phenomenon is based in a natural cycle of warming and cooling that has occurred for thousands of years. The crux of the climate change problem is that global warming has both known and unknown ramifications that impact individuals, regions, governments, and continents differently. The scale of the natural hazards that could be unleashed by continuing global warming trends does not fit current climate models. In spite of all our greatest efforts, predicting the intensity or location of the consequences of global warming is nearly impossible at this point.

In the past, most research has ignored the complex interrelationships among individual human groups, their environment, their social constructions, and the myriad different organisms and structures that interact with them each day. Because of this oversight, much information about climate changes and their impact has not been gathered. How does society address a problem whose scope is unfathomable at our current level of understanding? By addressing all known

aspects of the problem, implementing policy to mitigate all known impacts, and preparing for all unknown impacts simultaneously. Systemism may provide an important framework to address the climate change issue through its systematic and climatological aspects for 21st-century global public policy.

Conclusion

This chapter distinguishes two main methodological approaches, holism and individualism, and associates with them policy prescriptions of centralism and laissez-faire. Between these two perspectives, systemism offers a moderate way and common ground for reduction of political and social conflicts. Because of the nature of systems, the systemism approach may be more a mosaic than a single approach. It includes bits and pieces from a variety of disciplines. At the same time, interdisciplinary efforts can increase systemism's applicability to diverse fields. Systemism has influenced our current research methods in the areas of system analysis, case study, survey research, simulation, and sensitivity analysis. For political science in the 21st century, systemism will continue to contribute to our discussion of the political process, policy making, and management issues.

References and Further Readings

- Blakely, E. J. (2007). *Urban planning for climate change*. Cambridge, MA: Lincoln Institute of Land Policy.
- Bunge, M. (2000). Systemism: The alternative to individualism and holism. *Journal of Socio Economics*, 29, 147-157.
- Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.
- Collins, R. (1998). *The sociology of philosophies: A global theory of intellectual change*. Cambridge, MA: Harvard University Press.
- Denis, A. (2003). Methodology and policy in economic thought: A response to Mario Bunge. *Journal of Socio Economics*, 32, 219-226.
- Easton, D. (1965). *A systems analysis of political life*. New York: Wiley.
- Morgan, D. R., & Pelissero, J. P. (1980). Urban policy: Does political structure matter? *American Political Science Review*, 74(4), 999-1006.
- Pelissero, J. P. (Ed.). (2003). *Cities, politics, and policy*. Washington, DC: CQ Press.
- Starling, G. (2007). *Managing the public sector* (8th ed.). New York: Thomson Wadsworth.
- Williams, O. P., & Adrian, C. R. (1963). *Four cities: A study in comparative policy making*. Philadelphia: University of Pennsylvania Press.

5

RATIONALITY AND RATIONAL CHOICE

RICHARD A. WANDLING

Eastern Illinois University

The rationality concept has figured prominently in some of the most fascinating, heartfelt, and at times acrimonious scholarly exchanges among political scientists. This chapter focuses on five important intellectual developments in the study of rationality from a political science perspective: (1) the 1960s as an important era in scholarly exploration of the relationship between public policy making, decision making, and rationality; (2) Herbert Simon's seminal and hugely influential theorizing on decision making and the limits of individual rationality; (3) the legacy of bounded rationality, particularly in Graham Allison's models of decision making; (4) the seminal work of a group of economists and political scientists during the 1950s and 1960s who figured prominently in the emergence of modern rational choice theory; and (5) the modern scholarly debate over rational choice. A central theme of this survey is the tension between economic and political definitions of rationality and how these conceptions of rationality have shaped contemporary political science theory and research.

Policy Making, Decision Making, and Rationality

Charles Lindblom's "The Science of 'Muddling Through'" (1959) was an important milestone for a whole generation of theory and research on public policy making. Although

an economist by training, Lindblom became a major figure in political science, particularly among scholars of public administration and public policy. While exploring the intersection of public policy making and administrative decision making, Lindblom compares two "methods" of policy analysis and choice, identified as "rational-comprehensive" and "successive limited comparisons" (p. 81). The first method is summarized as the "root" method and the latter, the "branch" method. Lindblom presents the rational-comprehensive method (or model) in a negative light, as not only empirically flawed social science but as normatively questionable as a guide for sound decision making and public policy making in a democracy.

The rational-comprehensive model assumes that policies are crafted through a process that involves advance specification of key values and goals, tightly configured means-ends analysis, extensive analysis that is at once comprehensive and characterized by high levels of information, and a prominent role for theory-driven analysis. Out of this analytically intensive and information-rich process emerges a policy choice that is the "best" relative to decisional elements such as values and goals, actual analysis, and means evaluation. The successive limited comparisons model, however, is the one embraced by Lindblom. With this model, also known as incrementalism, values and goals often are not distinct, analysis of relations between ends and means is limited and perhaps even inappropriate, the options considered are few in number and

differ only marginally (or incrementally) from each other, and policy choices emerge out of a “succession of comparisons” (p. 81) among a limited set of options. If theory is important in the rational-comprehensive method, decision making in incrementalism is process oriented, with goodness of a decision defined as achieving agreement among analysts—that is, agreement rather than some objective evidence that the information, data, and analysis clearly point to the best option.

Lindblom’s framework represents a broadside against application of the rational model to policy making and administrative decision making. This comprehensively and tightly specified version of rationality does not work as either description or explanation of public policy making. However, to Lindblom this does not mean that policy making lacks rationality or is characterized by irrationality. It comes down to how rationality is conceptualized. Lindblom does not portray a chaotic or random universe with irrationality run rampant; there is a science or logic to “muddling through.” Decisions are made through a politicized process rather than based on compelling, objective logic of the facts, evidence, and information collected. In fact, to Lindblom the rationality of incrementalist-style policy making is preferable. Incrementalist-style rationality is very compatible with a pluralistic political system, particularly in producing options that rank high on political relevance and are grounded firmly in existing knowledge and information held by government officials.

Lindblom set the stage for further examination of rationality during the pivotal 1960s period of political science scholarship. Paul Diesing (1962) argued that rationality has multiple meanings and lamented the tendency to view rationality primarily as either technical or economic rationality concerned with organizational productivity and economic efficiency. Diesing develops a philosophy-oriented framework that argues for the study of three other forms of rationality—social, legal, and political. Aaron Wildavsky (1966), one of the 20th century’s most influential political scientists, takes the cue and warns strongly against framing rationality in terms of decision-making strategies or techniques such as cost–benefit analysis, systems analysis, and program budgeting. For political science, the latter were flawed because they indicated an economics-oriented view of rationality. To Wildavsky, political rationality is important in its own right because government leaders must calculate political costs such as the resources needed to generate support for a policy, the implications of a policy decision for reelection, and the possibility of provoking hostility for decisions not well received.

Simon, March, and the Limits of Rationality

Herbert Simon greatly influenced theory and research in fields as disparate as organization theory, decision sciences, and bureaucratic policy making. His ideas also

played a role in the development of rational choice theory—whether through his criticism or through efforts by some rational choice practitioners to incorporate Simon’s rationality assumptions into their research.

Simon provides a synthesizing approach to rationality that incorporates both economic and psychological dimensions while exploring the limits or boundaries of individual and organizational rationality. A starting point is Simon’s (1957) distinction between “objective” and “subjective” rationality. Objective rationality is evident if a decision or choice is the “correct behavior for maximizing given values in a given situation” (p. 76). With this version of rationality, a clear test is available to ascertain the correctness of a decision or choice. Subjective rationality incorporates psychological elements by considering the decision maker’s actual knowledge—or knowledge limitations. In short, based on the information possessed by the decision maker, what might be concluded about the rationality of a decision? Simon’s concern is that standards for achievement of objective rationality go well beyond the actual decision-making abilities of individuals, specifically individuals in complex organizations. The realities of psychology and human cognition mean that full knowledge of decision-related information is not possessed, and the full range of options also is not identified and evaluated in a comprehensive way.

Simon (1955) criticizes the rationality of classic economic theory and its model of “economic man” (p. 99), who is assumed to have extensive and intensive knowledge relative to the decision-making environment while possessing a well-organized and stable system of preferences, as well as a skill in computation that enables him to calculate the best alternative that reflects the highest point on his preference scale. The economic model of rationality is problematic for the development of a theory of the business firm or any type of organization, and this is the case whether the goal is empirically or normatively based theory. To Simon, real, empirical human rationality does not achieve the demanding standards of the classic economic model. Perhaps with a hint of things yet to come in the social sciences (including political science), Simon uses the term *rational choice* while inventorying key limits or constraints in “rational adaptation” behavior, particularly with respect to the range of alternatives considered, preferences, and decision maker knowledge of potential decision “payoffs” (p. 100). Simon also criticizes the “global rationality” assumptions that he sees embedded in game theory and castigates the economic rationality model as a “simplified model” that fails to capture the complex reality of a “choosing organization of limited knowledge and ability” (pp. 101, 114).

With this foundation, Simon fully develops his theory of *bounded rationality*—with important contributions from coauthor James March (March & Simon, 1958). The rationality of “administrative man” (p. 137) is compared and contrasted with the rationality requirements of classical economics—and statistical decision theory. In the

latter versions of rationality, decision optimality is the standard in an environment with a full and clear specification of alternatives, knowledge of consequences of the alternatives, and a “utility ordering” (p. 138) in which key values at stake guide fully conscious assessment of the alternatives. March and Simon, however, argue that individuals in organizational settings are not guided by the quest for optimality (i.e., the best possible decision) but rather make decisions at the point that an alternative is deemed satisfactory. They assert that “most human decision-making, whether individual or organizational, is concerned with the discovery and selection of satisfactory alternatives; only in exceptional cases is it concerned with the discovery and selection of optimal alternatives” (pp. 140–141). This point sets the stage for the much-referenced *satisficing* concept, which is a decision-making process in which the satisfactory standard is reached and the option selected is deemed as sufficient by the individual decision maker. In sum, the option is satisfactory—and it suffices. Satisficing is a major departure from the quest for the best possible choice as determined by extensively analyzing a wide range of alternatives and factoring in a full range of decision-related values or preferences. This model of decision making also parts company with the classic economic model in another way, through March and Simon’s assertion that alternatives are evaluated sequentially rather than simultaneously. At some point, an alternative is considered to be acceptable, given organizational goals, values, and decision-maker knowledge; the decision process concludes at that point.

Satisficing, however, does not take place in a vacuum; it is embedded in an organizational context in which rationality is bounded by “repertoires of action programs” (March & Simon, 1958, p. 169) that circumscribe and also channel the decision-making process. March and Simon give particular emphasis to the role of organization structure as the setting for individual decision making. Organization structure comes to play an important role in establishing the “boundaries of rationality” (p. 171). In essence, when we speak of the rationality of individual decision makers, we also are considering the role that organizations play in funneling or channeling decision making and even compensating for the limits of human rationality.

Later, Simon (1985) shed additional light on this path-breaking approach to rationality by noting that bounded rationality really is interchangeable with the term *procedural rationality*. Rationality is rooted in an organizational process of identifying alternatives, collecting information, and considering important values. This is another way of saying that there is no such thing as a substantively or objectively optimal decision. Simon sees this distinction as parallel to the concepts of procedural and substantive due process, observing that “in the same way, we can judge a person to be rational who uses a reasonable process for choosing; or, alternatively, we can judge a person to be rational who arrives at a reasonable choice” (1985, p. 294).

Bounded rationality is a way of focusing on the use of a reasonable process that helps to compensate for the limits of human rationality. And to avoid any misconceptions, Simon also contends that bounded rationality is not equivalent to irrationality. Objecting to the quality of choices or even the information that informed a decision is not the same as saying irrationality has prevailed. Individual decision makers do have goals and strive to make the best choices possible under the circumstances, which is another way of saying that they are “intendedly rational” (e.g., March & Simon, 1958, p. 170). Finally, Simon reminds us that bounded rationality has intellectual roots in psychological theory, specifically cognitive psychology. To Simon, cognitive psychology has a good appreciation of how individual choice making is limited in its computational abilities and involves a realistic understanding of individual problem-solving processes.

The Legacy of Bounded Rationality

The bounded rationality concept has figured prominently in political science, including influencing Lindblom’s incrementalist theory of rationality. Bounded rationality is a robust concept that lends itself readily to multiple meanings and applications, and it continues to play a role in how political scientists frame rationality. To illustrate, Jones (2003) evaluates the contributions of bounded rationality in public administration and public policy scholarship and argues that the bounded rationality approach has yielded an enhanced understanding of how government organizations may produce unexpected or even unpredicted policy or program results. With public organizations not operating under full rationality conditions, administrators aspiring toward rationality may nonetheless find their goals undermined by a variety of forces, such as informational uncertainties and nonrational elements of organizational decision making.

Bounded rationality also plays an important role in Allison’s (1971) three decision-making models for studying the Cuban missile crisis: rational policy, organizational process, and bureaucratic politics. The first and second models are most relevant to this chapter. Model 1 (rational policy) is Allison’s version of the economic rationality model, with assumptions of advance specification of goals and objectives; identification and evaluation of a range of options; clear-headed knowledge of consequences of decision alternatives, particularly with respect to costs and benefits; and finally selection of the best option from the standpoint of value maximization. This model conceptualizes decision making by the U.S. government as a unified national actor coolly mapping out a set of different alternatives for careful, deliberate evaluation—major options such as doing nothing, diplomatic pressures, a surgical air strike, or a blockade. Model 2 (organizational process) focuses on organizational processes and outputs, seeing U.S. decision making as the result of complex bureaucratic

properties. Simon's satisficing concept is evident in Allison's argument that decision making in Model 2 involves "sequential attention to goals" (p. 82). Bounded rationality also is evident in Allison's emphasis on "standard operating procedures" and "programs and repertoires" (p. 83) that coordinate the activities of individuals in government departments and agencies. These latter principles serve as the basis for Allison's much quoted examples of how organizational procedures and constraints may come to shape decision making at the highest levels of a presidential administration. Perhaps the most widely cited rationality example from Allison is Secretary of Defense Robert McNamara's argument for a political and internationally sensitive approach to blockade implementation, as opposed to admiral George Anderson's reluctance to deviate from the Navy's standard operating procedures for blockade placement.

Some scholars, however, have suggested that there may be problems with Allison's application of his decision-making models. To illustrate, Bendor and Hammond (1992) criticize Model 1 as unduly simplistic in its version of rational choice, and they contend that Allison has misinterpreted and misapplied bounded rationality theory. They argue that Allison's version of bounded rationality misinterprets Simon by viewing organizational structure, processes, and routines as a hindrance to quality decision making. Organizational properties such as standard operating procedures really are positive features in Simon's bounded rationality, by facilitating and assisting the decision-making process: In essence, complex challenges and difficult choices require that rationality be boosted through organizational processes, including processes as seemingly mundane as standard operating procedures. Organizations do not limit rationality; they facilitate rationality.

The Foundations of Rational Choice

The roots of modern rational choice theory generally are traced to the seminal contributions of a group of economists—primarily Arrow, Downs, Buchanan and Tullock, and Olson—and one path-breaking political scientist—Riker—through the 1950s to mid-1960s (e.g., see Almond, 1991; Ordeshook, 1990). Some scholars note the early formative role of social or economic philosophers such as Thomas Hobbes and Adam Smith (Monroe, 2001). Kenneth Arrow's (1963) social-choice approach to rationality is a good place to start. First developed in the early 1950s, it has contributed to decades of theory and research on the question of whether individual and collective rationality are inherently in conflict in democratic society. Individual rationality as indicated in expressed preferences might generate problematic collective social choices that lead to serious questioning of the possibility of coupling rationality with democracy—that is, without dictatorship to force choices on people. This puzzle is covered in

rational choice investigations of what generally is identified as the *possibility theorem*, or alternatively the *impossibility theorem*.

Anthony Downs's (1957) *Economic Theory of Democracy* is arguably the most important contribution from someone who is not a political scientist to rational choice in political science. While exploring the meanings of economic and political rationality, Downs presents a theory of rationality in which individuals in political and governmental arenas are guided by self-interest as they pursue choices with the highest levels of utility. The concept of utility figures prominently in economics and is a general way of summarizing the benefits choices bring to decision makers, and the utility concept makes regular appearances in the rational choice literature of political science. To Downs, benefits are not limited to a narrow monetary or financial nature; utility also may be derived from government services such as policing, water purification, and road repairs.

Downs is particularly well-known for his propositions on how self-interested voters assess the appeals of rationally oriented political parties in democratic political systems. These voters may also experience degrees of uncertainty and even information gaps, somewhat similar to what occurs in bounded rationality conditions. Kenneth Shepsle and Mark Bonchek (1997), coauthors of the standard text on rational choice, note the importance of Downs in spatial modeling to show how rational voters evaluate the merits of politicians and electoral candidates in ideological space. Governments themselves figure in Downs's analysis because government officials and political parties seek to maximize support from voters—for example, through spending on government programs or offering programs that appeal to voter self-interest. According to Downs (1957), governments are run by self-interested individuals whose primary concern is not an abstract ideal of social welfare maximization or the public interest; they are oriented toward developing government programs in relation to strategies to please voters.

James Buchanan and Gordon Tullock's (1962) *Calculus of Consent* presents a rationality model in which individuals choose according to the "more rather than less" principle (p. 18). The average individual seeks to maximize utility and secure more of what he or she values—rather than less of it—in the political arena as well as elsewhere. Buchanan and Tullock are particularly interested in the relationship between individual and collective rationality. Although they acknowledge that rationality in market-based decision making does not hold up as well in the governmental setting, they nonetheless argue for applying the logic of economic-based decision making to democratic political systems. Rational members of democratic society will decide in favor of political organizations and institutions that serve their respective individual interests, with competition among individuals also evident in this process. This competition becomes manifest as rational

individuals in constitutional democracies pursue more rather than less for themselves in the political arena. Although there may be some slippage from the full rationality standard regarding information levels of individuals and even the extent to which self-interest may dominate, Buchanan and Tullock confidently assert that “each participant in the political process tries, single-mindedly, to further his own interest, at the expense of others if this is necessary” (p. 305). Furthermore, individual choice plays out in an existing constitutional system—for example, the institutions, processes, and rules of representative democracy. In this sense, Buchanan and Tullock embrace a version of bounded rationality in that constitutional democracy also sets the boundaries for political choice.

Mancur Olson’s (1965) *Logic of Collective Action* represents a major challenge to traditional thinking on individual participation in groups in democratic society. Rational individuals may not have an incentive to join or participate in large voluntary associations, particularly those characterized as “latent” groups, if they can benefit from the collective or public goods provided by these groups without having to pay dues or incur other costs of membership (pp. 58–59). A key element of Olson’s approach to rationality concerns the “objectives” pursued by individuals. Olson pointedly makes the following observation:

The only requirement is that the behavior of individuals in large groups or organizations of the kind considered should generally be rational, in the sense that their objectives, whether selfish or unselfish, should be pursued by means that are efficient and effective for achieving these objectives. (p. 65)

Rational Choice Arrives in Political Science

William Riker’s (1962) *Theory of Political Coalitions* is probably the most important scholarly work in the emergence of rational choice in political science. Riker takes the theories of economics and mathematics-based game theory and expressly applies them to political decision making, presenting an alternative to political science’s long-standing focus on concepts such as power and authority. Riker sees rationality in terms of individuals who seek to win, rather than lose, in the context of various types of two-person games: “Politically rational man is the man who would rather win than lose, regardless of the particular stakes” (p. 22).

Whether considering topics such as voting choices or federal system design, Riker (1990) conceives of political rationality as involving actors who are “able to order their alternative goals, values, tastes, and strategies” and who “choose from available alternatives so as to maximize their satisfaction” (p. 172). In Riker we see the fusion of the rational actors of game theory and economics, transposed to the world of politics and government. Riker, however, sees his approach to rationality as transcending traditional arguments over pure economic and bounded rationality. The focus of rational choice theory should be on how

individuals decide with information available to them, from knowledge of their own preferences or through the consequences of alternatives themselves. His definition of rationality “requires only that, within the limits of available information about circumstances and consequences, actors choose so as to maximize their satisfaction” (p. 173). Riker became one of the most controversial figures in modern political science, arguing for political science to openly embrace rational choice as its future, particularly because “in contrast to economists, political scientists frequently have been methodically unsophisticated” (p. 178).

Riker’s approach to studying politics illustrates prominent features of modern rational choice. First, there is the common use of what may be called the “as if” assumption of rationality to guide empirical analysis and research (e.g., Moe, 1979). Individuals are assumed to act “as if” they decided according to principles such as utility maximization and the pursuit of self-interest (see Riker & Ordeshook, 1968), and then researchers go about the process of testing their propositions and hypotheses against empirical reality. The “as if” approach in rational choice theory has prompted great debate over rational choice’s approach to knowledge in the social sciences, with one writer exploring tensions between “instrumentalist empiricism” and “scientific realism” in rational choice scholarship while asking whether the “as if” assumption approach represents a “useful fiction” (MacDonald, 2003).

A second feature is the tendency of rational choice practitioners to work out anomalies or counterevidence from within the rational choice tradition itself—that is, to focus on what some refer to as the maintenance of core elements of the rational choice theory as a way of explaining political reality—even in the face of potentially confounding data or developments (e.g., Shapiro, 2005). To illustrate, Riker and Ordeshook (1968) addressed the puzzle that voting itself might be an irrational act when considering individual costs and benefits; they find that there really is an underlying rational calculus to the decision to vote—or for that matter not to vote.

A third feature of rational choice is its ongoing evolution, as we would expect of any healthy scholarly approach. The rational choice of recent decades is not the same as that of the 1960s and 1970s. In Riker, this is seen in his devotion in the latter part of his career to a scholarly approach labeled *heresthetics*, which focuses on the strategic use of communications, such as sentences and languages, by political leaders and elites in important arenas such as agenda control and coalition formation (Shepsle, 2003).

Understanding Contemporary Rational Choice Theory

Rational choice theory draws from the general approach called *rational actor theory*, which Monroe (1991) identifies as emphasizing individuals who pursue goals and decide among competing alternatives while possessing extensive information, a coherent preference ordering, and a commitment to the principles of self-interest and utility

maximization. Rational choice theorists, however, at times differ on how they incorporate these properties into their assumptions and empirical research. A major example is the distinction between *thin* and *thick* rationality. The thin version is the elemental approach to rationality that operates at a fairly broad level, not going much beyond general-purpose assumptions such as characterizing individuals as goal oriented, self-interested, and seeking utility maximization. A thickened version of rationality builds additional specifications into the rationality model—for example, actual belief systems, psychological needs, aspiration levels, cultural values, and even specific goals that may be important in the sociopolitical arena (e.g., see Ferejohn, 1991; Friedman, 1996). Rationality thus becomes richer or more substantive as it is thickened. The importance of understanding this distinction is underlined by Ostrom (2006), who criticizes the tendency in political science to “lump all scholars together who use a thin model of rationality together with those who are developing second- and third-generation behavioral theories” (p. 8).

A few examples from within rational choice scholarship illustrate efforts to broaden its framework and scholarly focus, particularly through the study of institutions. Shepsle and Barry Weingast (1994) assess the transition from the first generation of rational choice congressional research, which fused a behavioral orientation with a strong focus on majority cycles coupled with a relatively abstract notion of the legislature. The second and third generations of rational choice research on Congress, however, shifted toward incorporating institutional structure variables—such as committees, subcommittees, and their rules—along with parties and leadership in the postreform era. Terry Moe (2005) provides a critique from within rational choice that although supportive of the promise of rational choice for political science nonetheless calls for a much more substantial role for political power in rational choice and its study of institutions—in settings that range from the U.S. bureaucracy on through nation-to-nation interactions in international politics.

Richard Feiock (2007) develops a set of hypotheses on regional governance institutions based on what he identifies as a “second-generation model” that incorporates contextual factors that shape and underpin individuals as rational actors. A thin version of rationality is set aside, and contextual factors show how rationality may be bounded—and thus provide an example of integrating bounded rationality into modern rational choice. An excellent example of this synthesis is found in George Tsebilis (1990), who argues that rational choice has unique qualities in its ability to explain behavior of rational actors in the context of political and social institutions that establish the rules of the game in which individuals assess their options and seek utility maximization. Tsebilis’s embrace of a rational choice that is bounded by institutional setting is particularly interesting in view of his application of it to comparative political analysis.

To this point, rational choice has been presented in a summative way to introduce the reader to its roots and key

influences while providing some sense of its present concerns. It must be noted, however, that any survey of rational choice runs the risk of oversimplification, and the student may be wise to consider the statement by one well-known rational choice practitioner:

I suspect the only thing all RC [rational choice] people would agree upon is that their explanations presume that individuals behave purposively. Beyond that, every manner of disagreement theoretical, substantive, methodological can be found. RC is an approach, a general perspective, within which many different models can be located. (Fiorina, 1996, p. 87)

In addition, the undergraduate student with an interest in rationality will encounter multiple references to the public choice, social choice, and rational choice schools, and these terms often are used interchangeably—either accurately or inaccurately (e.g., Friedman, 1996; Monroe, 1991). Within political science, the term *public choice* certainly has definite connotations, primarily due to its association with a well-known political science couple, Elinor and Vincent Ostrom, whose unique and influential versions of rational choice theory and research have been identified by some as the *Bloomington school* (Mitchell, 1988). Illustrative of the sometimes tricky terrain, the term *public choice* may also represent a general ideological orientation to some political scientists who view public choice as having limited application to the discipline. These political scientists contend that public choice is too closely associated with a market-based model that ultimately sees politics and government as hindrances to individual and collective welfare. In sum, rational choice is a multifaceted subject with different schools of thought and even the potential for stirring some emotions.

Rational Choice Controversies

A full understanding of rational choice requires knowledge of controversies associated with this approach in the political science discipline. The decade of the 1990s represents a key turning point, with the emergence of open and occasionally heated debate over the value of rational choice to political science. This decade includes Donald Green and Ian Shapiro’s *Pathologies of Rational Choice Theory* (1994) and subsequent scholarly exchanges such as those in *The Rational Choice Controversy: Economic Models of Politics Reconsidered* (Friedman, 1996). A survey of some representative criticisms from this era captures the intensity of this debate:

- Gabriel Almond (1991) asserts that the economic model of rational choice neglects scholarship in disciplines such as sociology, psychology, and anthropology, and its assumptions of human rationality, with their emphasis on utility-maximizing behavior, produce a conception of human rationality that has no “substantive content” and is akin to the Scrabble blank tile that “can take on the value of any letter” (p. 49).

- Green and Shapiro (1994) skewer rational choice as fundamentally flawed, both theoretically and methodologically. Although noting that it has constructed sophisticated formal mathematical models, they contend that the value of rational choice to political science is undermined by a set of deep-seated social scientific pathologies—for example, its theory-driven research with little interest in solving real political questions or problems and its research results that “do little more than restate existing knowledge in rational choice terminology” (p. 6).

- Stephen Walt (1999) criticizes rational choice’s growing reliance on formal modeling, highly sophisticated mathematical analysis, and game theory applications, which he sees as not enhancing international security studies—with “rigor mortis” the more likely scholarly result than methodological “rigor.”

The rational choice debate carried over into the first decade of the 21st century, though the intensity level of the debate certainly has waned in recent years. The Perestroika movement, which borrowed its name from the reform era of the Soviet Union, probably was the most significant development in the rational choice debate of the past decade. The year 2001 witnessed a multipronged effort by a coalition of disenchanted political scientists to reform the American Political Science Association and redirect political science scholarship in general.

The Perestroikan critics of the political science establishment grouped rational choice with formal modeling and quantitatively oriented research as they made their case against a style of political science perceived as actually diminishing genuine knowledge of government, politics, and policy. Perhaps the most colorful statement to represent the emergent criticism of rational choice is the following call to arms:

William Riker was fond of saying that political science was a sinking ship, and rational choice theory was the only tugboat that might bring it to port. It is truer to say that Riker’s disciples have acted as pirates out to hijack political science to a rather barren island. Their piracy is doomed to fail. (Kasza, 2001, p. 599)

While the early fervor of the Perestroika heyday eventually dissipated, additional critiques of rational choice later emerged in an edited volume with the colorful title of *Perestroika! The Raucous Rebellion in Political Science* (Monroe, 2005). While rational choice was not by any means the sole object of attention of this volume, rational choice took its lumps from some high-profile political scientists such as Theodore Lowi and Samuel Beer.

Toward Reconciliation

It has not been all slings and arrows over the past decade. A case in point is the assessment provided by a

scholar with a well-established record of questioning rational choice and who also has argued for an alternative framework rooted in political psychology—*perspective theory*, which focuses on identity at the individual, group, and societal levels. Kristen Monroe (2001) argues that the discipline “has wasted too much time debating the merits of rational choice theory” and that it is time to focus more fully on asking “what we have learned that may be utilized in the next stage of constructing more realistic theories of political life” (pp. 165–166). Ostrom’s (2006) framing of the issue as “Rational Choice—An Evil Approach or a Theory Undergoing Change and Development?” (p. 8) also merits consideration. While embracing the value of rational choice as part of a diverse modern political science and certainly not seeing it as an evil approach, she nonetheless acknowledges, as a rational choice practitioner herself, that factionalism in today’s political science may have multiple sources, including rigid adherence to a narrow definition of rationality: “Some of the factionalism does stem from the arrogance of those who consider the continued use of a narrow model of human rationality the essential qualification for doing good social science” (p. 8).

Conclusion and Disciplinary Directions

The past 50-plus years have shown great interest by political scientists in the meaning and applications of rationality. Lindblom’s incrementalism ushered in an era of theory and research on the limits of rationality in crafting and choosing public policies, and Wildavsky expanded on incrementalist theory as he made the case for political rationality over economic rationality. Simon’s seminal theorizing contributed greatly to knowledge of the realities and parameters of rationality by arguing that there are limits on decision-makers’ abilities to acquire and process information and assess options. Rationality is circumscribed or limited, with bounded rationality the condition of individuals as they make important political, policy, and administrative choices. Starting with Riker, rational choice theory elevated the question of whether political actors—from voters on through institutional actors such as political parties, elected officials, government bureaucrats, or even nation-states—are motivated primarily by an economic-based sense of self-interest and utility maximization. Rational choice political scientists answered in the affirmative to this question as they drew from scholars such as Downs, Olson, and Buchanan and Tullock—all of whom cut their academic teeth in the economics discipline. With political scientists such as Riker and the Ostroms laying the foundations, rational choice would become an important force in the discipline.

Alternative conceptions of rationality have spurred debate among political scientists, including expressions of resistance to the notion that politics and government may be understood through the prism of an economics-oriented model of individual and organizational decision making

and behavior. Scholars such as Lindblom, Wildavsky, and even Allison questioned the value of seeing policy making and government decision making as tightly structured processes of high-end rationality. Critics of rational choice argued against a political science that reduced the political arena to self-interested, utility maximizing political actors who could be studied through heavily assumption-laden theories and methodologies that make extensive use of formal modeling. Rational choice practitioners, however, have defended their scholarly approach while asserting that rational choice is not a monolithic enterprise, with scholars marching in lockstep. In response to criticisms of early versions of a stripped-down rationality, known alternatively as thin rationality, second- and third-generation versions of rational choice have emerged to incorporate more nuanced and developed understandings of rationality in politics and government—such as adopting bounded rationality assumptions and paying attention to the impact of institutional or cultural variables such as legislative rules and traditions.

Although the dialogue over rational choice has been animated and sometimes heated, it ultimately has been beneficial to modern political science. From the multipronged criticisms of rational choice theory, methodology and research voiced by Green and Shapiro in the 1990s on through the sometimes heated debates of the Perestroika movement at the dawn of the new century, political science certainly has indicated a willingness to address fundamental issues and questions. For example, what drives or motivates individuals or government officials to action? Are they fundamentally self-interested? Or are they capable of placing the public interest over personal, economic-oriented calculations of benefit or utility? What of the impact of social-psychological factors such as emotions, values, and identity? Is the political arena best understood as a venue explained by the basic concepts and tools of economics? Just how much information can political actors handle when making a decision—such as whether to vote for a candidate, align with a political party or ideology, express support for a public policy, or evaluate the performance of government officials? All these intriguing questions figure in the study of rationality in political science, and they no doubt will continue to shape future generations of theory development and empirical research.

References and Further Readings

- Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.
- Almond, G. A. (1991). Rational choice theory and the social sciences. In K. R. Monroe (Ed.), *The economic approach to politics: A critical reassessment of the theory of rational action*. New York: HarperCollins.
- Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New Haven, CT: Yale University Press.
- Bendor, J., & Hammond, T. H. (1992). Rethinking Allison's models. *American Political Science Review*, 86, 301–322.
- Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent: Logical foundations of constitutional democracy*. Ann Arbor: University of Michigan Press.
- Diesing, P. (1962). *Reason in society: Five types of decisions and their social conditions*. Urbana: University of Illinois Press.
- Downs, A. (1957). *An economic theory of democracy*. Boston: Addison Wesley.
- Feiock, R. C. (2007). Rational choice and regional governance. *Journal of Urban Affairs*, 29, 47–63.
- Ferejohn, J. (1991). Rationality and interpretation: Parliamentary elections in early Stuart England. In K. R. Monroe (Ed.), *The economic approach to politics: A critical reassessment of the theory of rational action* (pp. 279–305). New York: HarperCollins.
- Fiorina, M. J. (1996). Rational choice, empirical contributions. In J. Friedman (Ed.), *The rational choice controversy: Economic models of politics reconsidered* (pp. 85–94). New Haven, CT: Yale University Press.
- Friedman, J. (1996). Introduction: Economic approaches to politics. In J. Friedman (Ed.), *The rational choice controversy: Economic models of politics reconsidered* (pp. 1–4). New Haven, CT: Yale University Press.
- Green, D. P., & Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science*. New Haven, CT: Yale University Press.
- Jones, B. D. (2003). Bounded rationality and political science: Lessons from public administration and public policy. *Journal of Public Administration Research and Theory*, 13, 395–412.
- Kasza, G. (2001). Perestroika: For an ecumenical science of politics. *PS: Political Science and Politics*, 34, 597–600.
- Lindblom, C. E. (1959). The science of “muddling through.” *Public Administration Review*, 19, 79–88.
- MacDonald, P. K. (2003). Useful fiction or miracle maker: The competing epistemological foundations of rational choice theory. *American Political Science Review*, 97, 551–565.
- March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.
- Mitchell, W. C. (1988). Virginia, Rochester and Bloomington: Twenty five years of public choice and political science. *Public Choice*, 56, 101–119.
- Moe, T. M. (1979). On the scientific status of rational models. *American Political Science Review*, 23, 215–243.
- Moe, T. M. (2005). Power and political institutions. *Perspectives on Politics*, 3, 215–233.
- Monroe, K. R. (1991). The theory of rational action: Its origins and usefulness for political science. In K. R. Monroe (Ed.), *The economic approach to politics: A critical reassessment of the theory of rational action* (pp. 1–23). New York: HarperCollins.
- Monroe, K. R. (2001). Paradigm shift: From rational choice to perspective. *International Political Science Review*, 22, 151–172.
- Monroe, K. R. (2005). *Perestroika! The raucous rebellion in political science*. New Haven, CT: Yale University Press.
- Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Ordeshook, P. C. (1990). The emerging discipline of political economy. In J. E. Alt & K. A. Shepsle (Eds.), *Perspectives on positive political economy* (pp. 9–30). Cambridge, UK: Cambridge University Press.

42 • GENERAL APPROACHES IN POLITICAL SCIENCE

- Ostrom, E. (2005). *Understanding institutional diversity*. Princeton, NJ: Princeton University Press.
- Ostrom, E. (2006). Converting threats into opportunities. *PS: Political Science & Politics*, 40, 3-12.
- Ostrom, E. (2007). Institutional rational choice: An assessment of the institutional analysis and development framework. In P. A. Sabatier (Ed.), *Theories of the policy process* (2nd ed., pp. 21-64). Boulder, CO: Westview Press.
- Riker, W. H. (1962). *The theory of political coalitions*. New Haven, CT: Yale University Press.
- Riker, W. H. (1990). Political science and rational choice. In J. E. Alt & K. A. Shepsle (Eds.), *Perspectives on positive political economy* (pp. 163-191). Cambridge, UK: Cambridge University Press.
- Riker, W. H., & Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review*, 62, 25-42.
- Shapiro, I. (2005). Problems, methods, and theories in the study of politics. In K. R. Monroe (Ed.), *Perestroika! The raucous rebellion in political science* (pp. 61-76). New Haven, CT: Yale University Press.
- Shepsle, K. A. (2003). Losers in politics (and how they sometimes become winners): Riker's heresthetic. *Perspectives on Politics*, 1, 307-315.
- Shepsle, K. A., & Bonchek, M. S. (1997). *Analyzing politics: Rationality, behavior and institutions*. New York: W. W. Norton.
- Shepsle, K. A., & Weingast, B. R. (1994). Positive theories of congressional institutions. *Legislative Studies Quarterly*, 19, 149-179.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.
- Simon, H. A. (1957). *Administrative behavior: A study of decision making processes in administrative organizations*. New York: Free Press.
- Simon, H. A. (1985). Human nature in politics: The dialogue with political science. *American Political Science Review*, 79, 293-304.
- Tsebilis, G. (1990). *Nested games: Rational choice in comparative politics*. Berkeley: University of California Press.
- Walt, S. M. (1999). Rigor or rigor mortis? Rational choice and security studies. *International Security Studies*, 23, 5-48.
- Wildavsky, A. (1966). The political economy of efficiency: Cost benefit analysis, systems analysis, and program budgeting. *Public Administration Review*, 26, 292-310.

Note: For building a foundation in rationality studies, the following sources will be particularly helpful to students: Downs, *An Economic Theory of Democracy* (1957); Monroe, *The Economic Approach to Politics: A Critical Reassessment of the Theory of Rational Action* (1991); Friedman, *The Rational Choice Controversy: Economic Models of Politics Reconsidered* (1996); and Shepsle and Bonchek, *Analyzing Politics: Rationality, Behavior and Institutions* (1997). Advanced students of rationality in political science should consider Ostrom's *Understanding Institutional Diversity* (2005), in which the 2009 Nobel Prize winner and public choice proponent grounds nuanced rational choice premises in social, economic, and political settings that include bounded rationality, the impacts of institutional rules and procedures, collective decision-making processes, game theoretic strategies, and community characteristics. A shorter and more focused version of this approach can be found in her 2007 survey article "Institutional Rational Choice," dealing with its use in public policy research. Ostrom's scholarly effort is part of what she calls the institutional analysis and development framework that already has been applied by a variety of scholars to real and challenging policy problems such as land use conflicts and common pool resource exhaustion problems (e.g., deforestation and soil erosion).

6

PRINCIPAL–AGENT THEORY

JOHN JAMES QUINN

Truman State University

Imagine the worst happens, and you are falsely accused of a crime in a city in which you are vacationing and in which you do not know anyone. You are arrested and advised of your rights to an attorney and one phone call. Which lawyer do you call? Which lawyer has the best training in the type of issues for which you were arrested? Which one has the best record on these matters? Which one can you afford? Will the one you pick work hard for you? Or will you choose one that will do only a minimal job, which could land you in prison? How can you make the best and most rational decision with this lack of information?

Few people in this situation would have enough information concerning which attorney has the appropriate training or which is the most dedicated to clients. However, any lawyer contacted would know these things about themselves. This asymmetry of (or differences in access to) information constitutes one of the key elements of the *principal–agent* problem. It is also called the *adverse selection* problem.¹ You have to hire someone with only limited information concerning his or her qualifications, training, and achievements. Moreover, potential agents have an incentive to overstate their abilities and experiences in order to obtain the commission.

In addition, you and your lawyer have some interests in this case that are different. You may want your lawyer to dedicate the next several months of his or her life to your case. Your lawyer, by contrast, may want to

spend the minimum amount of time possible on your case, get paid, win if possible, and pursue other interests. This highlights another aspect of the principal–agent problem: *moral hazard*. Agents and principals often have competing self-interests, despite the fact that the agent is hired specifically to represent the interests of the principal. The agent (the lawyer in this case) could put his or her interests ahead of the principal's (you in this case) by taking payment and not putting forth a strong effort. However, to be fair, unless the lawyer insists on payment up front, he or she runs the risk of doing a great job and then not being paid appropriately afterward.

Finally, your lawyer may recommend a course of action (e.g., plead innocent, take a plea bargain) that you cannot adequately analyze, given your inexperience with the law. You do not know the judges, the law, or past outcomes in similar cases. Lawyers should, and this is why they are hired. However, an incompetent lawyer will recommend a strategy with the same confidence as a well-qualified lawyer. This is also an aspect of the adverse selection problem. You would have to hire an additional lawyer or law firm to determine fully the quality of the strategy being employed. This asymmetry in knowledge and expertise between the people who could enter into a contract lies at the heart of the principal–agent relationship. (Otherwise, you could represent yourself in court and forgo hiring an agent to handle your business.)

Another common situation in which one can see this paradigm at work arises in the personal housing market. Imagine you want to buy a house in a new city and engage a real estate broker. This person has knowledge that you do not have (e.g., the quality of different neighborhoods, the quality of schools, how quickly homes sell, the quality of the contractors, the process of buying a house in that locality), and for this knowledge you hire the broker to become your agent. Typically, agents get some set percentage of the price of the house once it is sold. So when the agent discovers your top price, he or she may tend to show you the houses that are at the top, or slightly over, your price range. Why? The agent has a personal interest in selling you the most expensive house you can afford because he or she will make more money from that sale than from a house that sells for less; 6% of a higher number is more than 6% of a lower number. Your self-interest is in paying the least you possibly can for a house that gives you the most value. The agent may be better off (in a narrow sense) if you buy a house you cannot afford, as long as the deal goes through. If you have to sell your house because you cannot afford it, the agent loses nothing, and he or she might even be able to resell the house and earn money again. However, if the agent is concerned about reputation, he or she may strive to get you into an affordable house so that you pass on good reports to other potential clients.

That Person A would hire Person B to do the bidding of Person A and that their relationship is bound by a contract are well established in history. So too is the idea that the agent has some normative, legal, or moral obligations to do what is in the best interests of the principal. In legal terms, the agent has a *fiduciary obligation* to the principal. In fact, the principal–agent relationship is a key concept of British common law and in the field of law generally, where it is called *agency*. This idea of agency is therefore pervasive in society. According to Ross (1973), “All contractual arrangements, as between employer and employee or the state and the governed . . . contain elements of agency” (p. 134).

To understand the contemporary use of this approach in the social sciences, one must turn to political economy (Ordeshook, 1990). The principal–agent relationship, or agency, exists at the intersection of economics, political science, business, law, finance, and sociology. It is used as a heuristic tool to understand economic, social, business, or political relationships using self-interest as a guiding principal. This approach uses some, but not all, of the assumptions of microeconomics: that human interactions are best understood as a meeting of two self-interested and rational actors with relatively fixed preferences trying to maximize their own utility—but both of whom have less than complete information.

In political science, the principal–agent relationship is usually studied by rational choice scholars. The rational choice paradigm uses economic assumptions of human

nature to study political outcomes. As such, rational choice scholars begin with assumptions of rationality as well as the maximization of (relatively) fixed goals. These are the *strong* assumptions of rational choice. For example, the assumption of wealth maximization often translates to power maximization or reelection for political leaders (Levi, 1997). It also includes several weaker assumptions, including no information costs; no transaction costs; no collective or organizational costs; no transportation costs; and no role for history, institutions, or culture. There are simplifying assumptions that are not true, per se, but they are held to be true for the parsimony of the model. However, some authors do not include all of the assumptions (or they lift or “assume away” one assumption or another) and examine the likely outcomes of no longer having all the simplifying assumptions in the model. However, different scholars have examined political interactions and have lifted one assumption or another. Olson (1965), for example, lifted the assumption of collective action costs to show how by reintroducing these costs, one could predict more realistic political outcomes than before.

Therefore, the principal–agent paradigm is used to describe situations in which information costs, which are normally lifted or assumed away in microeconomics, are reintroduced, along with risk sharing or risk shifting. In the same way that physicists assume away friction to describe “ideal physics,” so do economists sometimes assume away these other costs of exchange. However, the principal–agent paradigm retains the strong rational choice assumptions of self-interest on the parts of both the principal and the agent, as well as their relatively fixed preferences. The inclusion of transaction costs results in both *adverse selection* and *moral hazard* (Moe, 1984).

Let us explore in more depth the key elements of adverse selection and moral hazard in the principal–agent paradigm. An illustration of adverse selection may be seen when applicants for a position claim greater skills or work habits than they actually have, believing that these personal attributes are difficult for potential employers to know. Prospective employees (agents) have the incentive to exaggerate their abilities (or pad their résumé) to get their foot in the door, and then they quickly learn the skills after being hired. The firm doing the hiring (the principal) wants to hire someone already qualified and who brings these skills along. However, the principal can determine only imperfectly whether a prospective employee actually has these attributes or is exaggerating.

The moral hazard aspect of the principal–agent relationship usually occurs after a person is hired and when the amount or quality of work performed (or output) is difficult to measure or monitor. For example, if a contractor (i.e., agent) is building a house for you in your town, you can stop by the site and witness the various parts of the house being constructed. If, however, you hire an agent to manage a plant in a distant state, you would have to spend

more time and resources to visit that factory. Moreover, if the agent knows when you are to arrive, he or she can look busy during the visit and hide low productivity.

In addition, even with supervision, sometimes output is inherently hard to measure. For example, if a teacher has many students who do poorly on performance tests, who should be held accountable? Should one blame the teacher who was unable to motivate otherwise “good” students? Or should one blame the “underachieving” students who did not appreciate the high quality of the education they were receiving? In addition, should pay be linked to student outcomes? Similarly, should a car salesperson with a slow month be blamed and seen as a slacker? Or could the salesperson have been quite diligent, but the market turned against him or her that month?

Classic Views of the Principal–Agent Relationship

According to Miller (2005), the canonical or orthodox principal–agent relationship is marked by several key ideas. First, the actions of the agent affect the wealth or well-being of the principal, so the principal is expecting some payoff (reward or punishment) arising from actions of the agent. Second, information asymmetries exist, and the agent has special knowledge or abilities, the lack of which keeps the principal from doing the job himself or herself. In addition, the principal can learn of the efforts of the agent only with difficulty or at a high cost. Third, the principal and agent are assumed to have some preferences that differ. As mentioned above, the agent usually wants the most pay for the least amount of work, and the principal wants the biggest reward (or smallest penalty) with the least payment to the agent. The employee–employer relationship provides a classic example. Another aspect of different preferences is that the agent is usually risk averse whereas the principal is considered risk neutral. To understand this, imagine someone offered two different jobs selling cars, one that pays purely on commission and the other that pays a base rate plus a much smaller commission. Even if the expected income were higher at the commission-only job, the risk-averse agent would likely pick the job with the base pay. Risk-neutral people would maximize their income by taking the commission job because its potential for earning money is greater.

Fourth, the initiative for creating a contract lies with a unified (single) principal. When the principal engages the agent, the principal sets the terms of the contract, to which the agent agrees or not. Fifth, both agent and principal know the basics of the process in which the agent will be engaged. Although a principal may not be able to monitor directly how much work an agent does, the principal is assumed to know the results of the agent’s work, which can be used to infer the efforts of the agent. This is called *backwards induction*—assuming high efforts on the basis

of good outcomes and low efforts on the basis of poor outcomes. Information about the shape of the game and the outcomes of the work of the agent are inexpensive to obtain. Last, the principal can set the specifics of the contract, which include the incentive structure for the agent. The principal, using logic, can determine the best contract to obtain the most wealth from the agent, given the first five elements of the relationship between principal and agent.

Therefore, the principal–agent dilemma exists when a principal wants to hire an agent, but given the self-interest of the agent, as well as the agent’s unknown qualifications, the principal’s choice could lead to poor performance on the agent’s part, thereby harming the principal. The principal therefore tries to create a contract, or payment system, such that the agent acts on the principal’s behalf. In other words, their economic interests are aligned through a contract that shapes incentives.

As a result of the above elements of the principal–agent model, the rule of thumb for the relationship is that where the costs of monitoring the agent are high (and the agent is risk neutral), the principal creates a contract that links payment chiefly to outcomes rather than making trips to monitor the agent or devising complex systems of accountability. For example, in the case of civil law, plaintiff lawyers are often paid with contingency fees and only when the outcome is a successful verdict. If the lawyer (agent) does not win the case, then he or she is not paid. For many in sales, some commission system is usually found, often on top of a small salary, given the risk-averse aspect of the agent. Finally, CEOs are often paid some mix of salary and stock bonuses that align their interests with the price of stock shares; in other words, the agent’s interests become aligned with the owners’ (stockholders) interests in higher stock prices.

Where it is relatively easy or low in cost to monitor the actions of the agent (or where the agent is particularly risk averse), then payment is tied to efforts and not outcomes. In fact, many people are paid by the hour because their employers share the view that time at work results in desired outcomes. For example, a mechanic or carpenter who hires an assistant can easily monitor the efforts and quality of work put forth by the assistant. Therefore, the mechanic or carpenter can pay the assistant by the hour without significant adverse selection and moral hazard effects.

Much of the literature on principal–agent theory examines the types of contracts that are optimal under different levels of risk and transaction costs. The dependent variable is the optimal or second-best contract designed by the principal and imposed on the agent. Many authors explore how transaction costs affect the type and nature of the contract created. Much of this literature uses complex mathematical models and proofs to arrive at conclusions.

A leitmotif running through these ideas is that a contract creates incentives that can align the interests of the agent

with those of the principal. Within economics, the principal–agent approach originated in the context of insurance and credit markets and examinations of how principals can extract rents from agents (Stiglitz, 1987). Since it is assumed that the principal has control over the contract, a rent-extracting agent is not usually modeled.

According to Jensen (1983), there are two basic strands of principal–agent theory. The first is a more empirical, or positive, approach, often called the *positive theory of agency*. This approach attempts to arrive at testable hypotheses derived from the particular assumptions of the model (such as the six described above). This approach is more often seen in political science. The positive theory of agency can also be used as a case-illustrative approach to explain phenomena unexplainable by other paradigms or models. The second strand of this literature is economic in orientation, with formal modeling and advanced mathematics used to make predictions about which contracts would be the most efficient under various conditions. It is often called, *tout court*, principal–agent. It is highly formalistic and mathematical, and rarely empirical.

However, especially on the agency side of the approach, one understands that contracts made between agents and principals occur in a broader societal context than a one-time negotiation resulting in a contract. Government institutions, or regulations, create background conditions in which negotiations or contracts are constructed, especially through the regulation of socially recognized agents. For example, each state in the United States regulates lawyers and empowers the state bar to ensure that each lawyer has met minimum qualifications. If, indeed, determining the quality of any particular lawyer involved no information costs, then making lawyers pass the bar examination would be needless—one would know ahead of time which lawyers were prepared and which were not. But since information is costly, and the potentially harmful societal effect of poorly trained lawyers would be great, the government regulates this market. With the requirement of passing the bar, the chances that a lawyer will have little or no training are much lower. These sorts of regulations establish a floor, rather than a ceiling, for these interchanges. Similarly, medical doctors have to pass exams to practice medicine. Finally, teachers are certified in nearly every state for employment in the public schools: They must take a minimum number of hours in education courses from a credentialed school of higher learning. Each state's requiring potential agents to have met minimum qualifications reduces, but does not eliminate, adverse selection problems facing potential principals. Government can also regulate food quality, water quality, and more to ensure the smooth working of markets under conditions of transaction and information costs.

Similarly, but expanding beyond the narrow principal–agent problem, both government and private institutions can be seen as forms of megacontracts, or the guidelines under which other contracts are written. Some argue that

institutions arise precisely because of information and transaction costs (which give rise to the principal–agent dilemma). As discussed above, the solution to the principal–agent problem is found in the contract itself. Some suggest that most organizations “serve as a nexus for a set of contracting relationships among individuals” (Jensen & Meckling, 1976, p. 310).

One of the pioneers in transaction costs, Ronald Coase (1937), held that the rise of the contemporary business firm—which dominates the landscape of U.S. business—makes sense only when one accounts for transaction and information costs. The existence of these firms would be illogical if these costs were truly zero. Why have an economic institution dedicated to the production of various elements for some product when one could costlessly contract out all the subcomponents? Given the lack of complete information on quality, performance, punctuality, and price for each subcontractor, adverse selection and moral hazard necessarily arise, and these problems have been solved by the evolution of production within a firm where relations were based on hierarchy and authority instead of only price.

As such, formal hierarchical organizations can be used to improve monitoring to reduce moral hazard and create incentive structures, or contracts, such that agents perform the duties expected of them by the principals. Where monitoring is difficult, the contract can employ “proxy” measures of actual quality instead of quality itself (Moe, 1984). Some examples are hours worked, products produced, and sales completed. However, where these things are difficult or costly to measure, others are hired to monitor the work of employees. Therefore, supervisors in stores who are not stocking shelves or sweeping floors are there to make sure that those employed to do so, do so. Lower-level management in firms is hired to monitor employee performance. And, in turn, midlevel management monitors lower-level managers.

Similarly, Williamson (1985) examined economic institutions to see how they form economic incentives that shape human behavior. In fact, Williamson suggests that the principal–agent relationship is a subcomponent—along with property rights (e.g., North, 1981)—of the literature on transaction costs, which looks at incentive structures to explain outcomes. The other branch of the literature on transaction costs examines institutions that rely on non-market governance and measurement to have economic outcomes. Williamson claims that traditional microeconomic theory falsely views all deviations from “pure market behavior” to be caused by monopoly. Ironically, economic institutions reduce the very transaction costs (prices and premiums for risk, uncertainty, bargaining, and obtaining information) that are assumed away in traditional microeconomic theory. Jensen and Meckling (1976) suggest that organizations themselves are actually multi-lateral contracts between and among many principals and agents—thus bridging the principal–agent literature with

neoinstitutional ideas. Therefore, the incentives found in contracts are also found in institutions.

Although not all principal–agent relations examine institutions and their roles, one can see how, empirically, society has helped to overcome information costs and make markets operate “as if” information costs were low or nonexistent. Institutions can also play a role in shaping incentives between principals and agents, which act as a form of contract. Therefore, contracts can be seen as more concise versions of institutions, and institutions are longer established rules and norms that create incentives—which in turn shape the interchange between principal and agent.

In addition, private companies can help reduce uncertainty and thereby lower information and transaction costs for consumers if doing so helps private companies move product. Imagine that a new car company with no record of accomplishment wants to enter the market. In order to reduce the risk to the customer, the car company can give a warranty that the car will perform well for X number of miles, or X number of years. So even if the car proves to have many mechanical problems, the customer is somewhat protected from the lack of information before the purchase of the car. This should lower customers’ risk of purchasing a new product that does not yet have a record of accomplishment (and thus less information about quality).

Firms also realize that potential customers incur information costs in making decisions. For example, firms that win awards for quality or safety will highlight this in their advertisements as a form of external verification for otherwise ubiquitous claims of quality and performance. Knowledge of the awards reduces the risk, or uncertainty, the client faces, because uncertainty results from gaps in information. Some information, such as the life span of any particular car, is unknown to both seller and buyer, although the car salesperson has a better probabilistic understanding of how well the different models of cars he or she sells will perform.

Political Science

In early political science, Max Weber identified a construct similar to the principal–agent perspective when he discussed the state and sources of authority. He argued that there are three key constituencies in a state: (1) the power holders, (2) their servants (the bureaucracy), and (3) the population (Lane, 2008; Weber, 1978). One can apply the principal–agent relationship to the state and think of the population as the principal and the power holders as the agents (Lane, 2008). Alternatively, one could think of the power holders as the principals who have to monitor the bureaucracy, who are the agents. An example would be congressional oversight of bureaucracies (Weingast, 1984).

However, some might argue that thinking of the population of a country as the principals and officeholders as the agents could be problematic for several reasons. First,

do all the principals have a unified interest that the agents can understand? In fact, no matter what politicians do, one could argue that they are responding to some societal demands while ignoring others. Therefore, the lack of a clear and unified principal could make the application of this paradigm to politics more difficult. Moe (1984) discusses the difficulty of analysis with multiple principals.

One possible exception to this problem may be the case of corruption. Several scholars have used the principal–agent model to describe corruption (Alam, 1989; Klitgaard, 1988; Quinn, 2008). Since corruption can be difficult to define, Alam (1989) argued that corruption can be best understood as a function of all principal–agent relationships and as such may be defined as “(1) the sacrifice of the principal’s interest for the agent’s, or (2) the violation of norms defining the agent’s behavior” (p. 442). Here, the principals are the people, and the government is the agent. The normal assumptions are that the population would have a united interest in the most public goods at the lowest costs, and corruption would eat away at either the quality or the quantity of the service rendered. Quinn (2008) suggests that when the political elite becomes the agent for the people (principals) through majority state ownership of most capital-intensive industries or the largest export sector, then the principal–agent problem predicts a rise in corruption. This would be especially true within these economic sectors—to the point that the potentially most productive sectors of the economy can become a drain on wealth instead of an engine of growth. The extreme form of betrayal of public interests by the ruling class (agents) could be seen as predatory rule (Levi, 1988).

Some political scientists studying Africa have argued that politicians could be acting “rationally” when they are following “irrational” economic policies. To understand this point of view, one must suspend the assumptions of both no collective action costs and no transaction costs. Since the agents (the politicians) are maximizing their political power (i.e., incumbency and power) and responding to the part of the society that is best organized (e.g., other elites, urban dwellers, the military), they could establish economic policies that provide a return only or primarily to powerful segments while claiming to pursue the public interest. These powerful actors will be parts of society that can overcome their collective action costs. Thus economic policies that enrich a small segment of politically powerful people are enacted, even though they lead to economic performance that undermines the interests of most of the principals (the majority of citizens; see, e.g., Bates, 1981). (Although few such analyses are specifically pitched in principal–agent terms, they can be understood as such.) Again, some have argued that this betrayal of the principals by the agents is greater with majority state ownership of most economic assets (Quinn, 2002, 2008) because the normal economic interests that would lobby government for better policy are no longer separated from government itself.

Weingast (1984) applied this theory to bureaucracies, in which case the agents are the bureaucrats, and the principal is Congress (acting on behalf of the ultimate principals, the voters). Traditional analysis of how tightly Congress monitored bureaucracies found that the monitoring was very loose (Moe, 1984). However, consistent with the literature on principal–agent relations, when costs for monitoring the behavior of agents are high, then agents are held to account through outcomes. Since what members of Congress care about is reelection, if they are regularly reelected, then the bureaucracies are monitored by the desired outcome. Since members of Congress are reelected at very high levels, the need for more monitoring is low—this conclusion is quite consistent with the literature.² Therefore, if benefits from these bureaucracies flow to constituents’ needs sufficiently to allow reelection of the members of Congress, then the bureaucracies are not as autonomous from Congress as the loose monitoring implies.

Levi (1988) uses transaction costs, principal–agent relations, and discount rates to predict the types of tax systems different countries will use. She uses the assumption of revenue maximization on behalf of rulers who use agents (which have different costs to monitor) and the citizenry to show how each predicts different tax schemes. Different systems adopt different revenue collection schemes, and she illustrates how each was a response to different discount rates, transactions costs, and the self-interest of both principals and agents.

Lane (2008) argues that nearly all aspects of politics (i.e., the rise of states, political parties, levels of development) can be usefully seen through a principal–agent paradigm. Lane holds that the only way to fully constrain agents to do the bidding of principals is through the rule of law.

Critiques

Many critique this approach in a similar fashion as they would rational choice. First, some argue that people do not maximize their returns, *per se*; rather, the view of bounded rationality suggests that people “satisfice” instead (March & Simon, 1958; Simon, 1947). The expression “It’s good enough for government work” sums up this sentiment. Therefore, instead of searching for the most efficient contract between the principal and agent, this view holds that the agent does the minimum to satisfy the principal, and the principal does not work for the best contract but only for an acceptable one. However, with the role of information costs introduced into this paradigm, these two approaches are not necessarily at odds empirically, though they would be theoretically. According to the principal–agent paradigm, should the principal learn that the agent is satisficing instead of maximizing, then the principal would likely cancel the contract or seek another agent. However, with high monitoring costs, agents are held to outcomes, and if the outcomes were acceptable, then the relationship would have “worked.”

In addition, one could criticize this approach because it assumes (a) that human behavior is primarily motivated by economic gain or loss and (b) that each party is only pursuing self-interest. Engaging family members as agents could align goals of principal and agent more strongly than could pure economic motivation. Many African dictators and early European monarchs made sure to have family members in charge of the military or other key parts of the bureaucracy.

However, even loyalty for family members could be seen from a self-interested perspective. If a dictator is from a minority ethnic (tribal) group, then he or she often gives disproportionate access to top positions to members from his or her ethnic group or family. This behavior is logical because these agents have self-interest to be loyal to the ruler: They understand that their control over these jobs and resources would likely disappear should a new leader, especially one from a different ethnic group, replace the current one. This arrangement greatly reduces the agents’ returns under alternative arrangements, and their estimated returns under alternative ruling institutions would likely be quite low (see Quinn, 2002, Chapter 9).

Others may hold that the assumptions of principal and agent are too simplistic and that a bureaucracy (especially in the U.S. case) has several competing principals exercising authority over the agent. Those who adhere to the principal–agent paradigm would say that this caveat would not destroy the paradigm but would merely predict more autonomy for agents within the paradigm because principals would be aware that the agents have other principals to whom they are held accountable.

Finally, the formal modeling side of the principal–agent paradigm has been criticized as being too formalistic, with few real-life applications. Real salary schemes seldom match the theoretical models developed. Some have also suggested that the approach is tautological. However, if one uses a paradigm to explain behavior that was previously unexplainable, then the paradigm is useful. The same is true if testable implications of the paradigm emerge. The paradigm is not directly tested, but its ability to predict is (Jensen, 1983).

Possible Application to Recent Events

One can use a principal–agent approach to illustrate part of the recent international financial problem with “toxic” mortgage debts. At the heart of this issue is the fact that many real estate agents and banks sold houses to clients who ultimately could not afford them. Why? Several important elements were in place to shift risk and information. First, investors had limited information about the risk involved in real estate. For quite a while, housing prices seemed to be impervious to price declines. This “limited” or false information shifted incentives in the market as to how safe people felt real estate investments were. Second, financial instruments (called *derivatives* or *credit default swaps*) were created to reduce risk, but those

instruments containing housing mortgages were given quite low risk ratings when they were, in fact, high-risk instruments. Third, these investment tools were effectively forms of insurance and were not regulated by government, which made information costs about them higher. Since these instruments were effectively insurance, but not regulated as such, the banks lent more money against these assets than would have been allowed in regular insurance markets. This led to “overleveraging.” So instead of lending against assets at 3, 5, or even 10 times their value, the loans were leveraged up to 30 times. Therefore, the effect of each bad debt was multiplied 30 times in the system instead of only 3 to 10 times.

Nevertheless, why did the real estate agents and bankers sell houses to people who could not afford them or who could barely afford them? These agents had no incentives to be careful. The real estate agents, being on commission, were paid only when houses were sold. Also, it is important to note, they were not punished financially for putting people into unaffordable houses as long as the sale went through. They had personal incentives that were partially at odds with the principals with whom they did business.

The risk-averse part of the “normal” process of selling mortgages should have been at the level of mortgage banks. In the past, if people bought homes that went into default, the bank lost money. Therefore, it made financial sense for mortgage banks to pay the information cost to scrutinize borrowers’ abilities to repay these loans. However, since bankers knew their mortgages were to be repackaged and sold as mortgage derivatives and securities, the banks were able to move the risk on to others while locking in their profits. Since there was a widespread belief that real estate prices would not fall, others bought these mortgage-backed securities as “safe” investments. Then very high side bets (derivatives) were made on what should have been safe investments but which turned out actually to be risky bets. (Here again, transaction costs and information costs clearly impacted markets in a negative way, although these costs are not considered in microeconomics.) It is important that those engaged in the selling of homes and primary mortgages shared none of the risk and made lots of money shifting the risk downstream, so they sold as many homes as possible. This circumstance illustrates the moral hazard involved in the issue because some profits were divorced from the risks of their practices. Now the people who bought the risky mortgage-backed debts faced adverse selection because these loans were rated as nonrisky investments when they were, in fact, quite risky. Although a full analysis of the recent financial problems is beyond the scope of this chapter, one can see in basic relief some of the key elements of the principal-agent problem. One also sees that the microeconomic assumptions of complete information predicted perfect market behavior (i.e., the market is always right), which proved to be illusory and almost brought down the financial system.

The debate over merit pay for teachers is another current issue that could benefit from a principal-agent explanation. According to Holmstrom and Milgrom (1991), in cases in which the agents perform multiple and complex tasks that are not easily reducible to simple proxy measures, fixed-pay schemes align incentives better than outcome-based pay can. Should outcome-based pay (or merit pay) or promotions be adopted when agents have to achieve several competing goals, and only a few of those goals are directly measurable, then outcome-based rewards would incentivize agents to shift from the hard-to-measure aspects of their jobs to aspects that are easily measured. This is what many critics of “No Child Left Behind” meant when they said teachers would “only teach to the test.”²³ That is to say that the other high-quality items normally taught or required in schools (e.g., discipline, abstract thinking, homework, socialization, effective writing, creative writing, affective learning) were replaced with materials that were to be on the test (e.g., reading comprehension, vocabulary, basic math, geometry). Supporters of teacher merit pay suggest that these incentives would increase teacher effort, especially in poorly performing schools. The idea of raising all salaries to reduce average adverse selection may join these two arguments, especially at the worst performing schools.

In sum, the principal-agent paradigm, which is often used in economics, business, and political science, is a powerful tool, especially when the assumption that people act primarily out of self-interest is most appropriate. Although many argue about whether people maximize returns or look only for acceptable returns, this paradigm can illustrate how by adding the variable of cost of information to models, one can sometimes predict outcomes more accurately. It can also explain how institutions aid or hinder the process of holding agents to account or getting them to do the principal’s bidding when agents’ behavior is otherwise costly to monitor. It is likely a rising paradigm in political science, especially where rational choice (which emphasizes purposeful, self-interested behavior) meets institutionalism (which acknowledges that information asymmetries are real and costly; see, e.g., Hall, 1997).

Notes

1. This review is not meant as a complete review of the all studies on this issue. Rather, it intends to open the door of this literature for undergraduates. Citations emphasize early works, classics, illustrations of a few well known works, and recent reviews, not necessarily a representative sample of the most recent literature and its findings. Most readings that require calculus to understand have been ignored, aside from some general findings.

2. Scandals often result in hearings that appear to be monitoring the bureaucracy for high profile, symbolic issues.

3. In fact, reports of teachers cheating to help students pass the exams are rising. For example, see Axtman (2005).

References and Further Readings

- Alam, M. S. (1989). Anatomy of corruption: An approach to the political economy of underdevelopment. *American Journal of Economics & Sociology*, 48(4), 441-471.
- Axtman, K. (2005, January 11). When tests' cheaters are the teachers. *Christian Science Monitor*. Retrieved April 2, 2010, from <http://www.csmonitor.com/2005/0111/p01s03-ussc.html>
- Bates, R. H. (1981). *Markets and states in tropical Africa: The political basis of agricultural policies*. Los Angeles: University of California Press.
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4, 386-405.
- Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57-74.
- Geddes, B. (2003). *Paradigms and sand castles: Theory building and research design in comparative politics*. Ann Arbor: University of Michigan Press.
- Hall, P. A. (1997). The role of interests, institutions, and ideas in the comparative political economy of advanced industrial nations. In M. I. Lichbach & A. S. Zuckerman (Eds.), *Comparative politics: Rationality, culture, and structure* (pp. 174-208). Cambridge, UK: Cambridge University Press.
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics & Organization*, 7, 24-52.
- Jensen, M. C. (1983). Organization theory and methodology. *Accounting Review*, 63(2), 319-339.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3, 305-360.
- Kiser, E. (1999). Comparing varieties of agency theory in economics, political science, and sociology: An illustration from state policy implementation. *Sociological Theory*, 17(2), 146-170.
- Klitgaard, R. (1988). *Controlling corruption*. Berkeley: University of California Press.
- Laffont, J. J. (Ed.). (2003). *The principal agent model: The economic theory of incentives*. Northampton, MA: Elgar.
- Lane, J. E. (2008). *Comparative politics: The principal agent perspective*. New York: Routledge.
- Levi, M. (1988). *Of rule and revenue*. Los Angeles: University of California Press.
- Levi, M. (1997). A model, a method, and a map: Rational choice in comparative and historical analysis. In M. I. Lichbach & A. S. Zuckerman (Eds.), *Comparative politics: Rationality, culture, and structure* (pp. 19-41). Cambridge, UK: Cambridge University Press.
- March, J. C., & Simon, H. (1958). *Organizations*. New York: Wiley.
- Miller, G. J. (2005). The political evolution of principal agent models. *Annual Review of Political Science*, 8, 203-225.
- Moe, T. M. (1984). The new economics of organization. *American Journal of Political Science*, 28(4), 739-777.
- North, D. C. (1981). *Structure and change in economic history*. New York: Norton.
- Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Ordeshook, P. C. (1990). The emerging discipline of political economy. In J. E. Alt & K. A. Shepsle (Eds.), *Perspectives in Positive Political Economy* (pp. 9-30). Cambridge, UK: Cambridge University Press.
- Quinn, J. J. (2002). *The road oft traveled: Development policies and majority state ownership of industry in Africa*. Westport, CT: Praeger.
- Quinn, J. J. (2008). The effects of majority state ownership of significant economic sectors on corruption: A cross regional comparison. *International Interactions*, 34(1), 84-128.
- Rockman, B. A. (2001). Theory and inference in the study of bureaucracy: Micro and neoinstitutionalist foundations of choice. *Journal of Public Administration Research & Theory*, 11(1), 3-28.
- Ross, S. A. (1973). The economic theory of agency: The principal's problems. *American Economic Review*, 62, 134-139.
- Shapiro, S. P. (2005). Agency theory. *Annual Review of Sociology*, 31, 263-284.
- Shavell, S. (1979). Risk sharing and incentives in the principal and agent relationships. *Bell Journal of Economics*, 10, 55-73.
- Simon, H. A. (1947). *Administrative behavior*. New York: Macmillan.
- Spence, M., & Zeckhauser, R. (1971). Insurance, information and individual action. *American Economic Review*, 61(2), 380-387.
- Stiglitz, J. E. (1987). Principal agent. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Allocation, information, & markets* (pp. 241-253). New York: W. W. Norton.
- Weber, M. (1978). *Economy and society: An outline of interpretive sociology I II*. Los Angeles: University of California Press.
- Weingast, B. R. (1984). The congressional bureaucratic system: A principal agent perspective (with applications to the SEC). *Public Choice*, 44(1), 147-191.
- Williamson, O. E. (1985). *The economic institutions of capital ism*. New York: Free Press.

7

POLITICAL PSYCHOLOGY

DAVID PATRICK HOUGHTON

University of Central Florida

The term *political psychology* refers to the study of the ways in which human psychology—our thought processes, personalities, beliefs, and so on—affects politics, and it can be thought of as the area where the academic disciplines of political science and psychology overlap or intersect. It can also be thought of as a kind of “bridge” between the two fields. Just as political economy studies the ways in which economic relationships affect political behavior (as well as the ways in which politics affects economics), political psychology looks at the ways in which our cognitions and emotions, as well as the social pressures surrounding us, can shape our behavior in the political realm. It would be odd indeed if the ways in which the human mind works, for instance, did not affect our voting choices in significant ways, the manner in which we campaign, the tendency of some individuals to engage in genocidal behavior, or the practice of terrorism (to note but a few of the ways in which human beings act politically). In fact, while many political scientists attempt to explain our behavior in other ways—most commonly, by modeling it according to the assumptions of classical economics—there is at least a grudging acceptance within the discipline today that any full account of the vast array of behaviors that human beings engage in when they act politically simply *requires* an understanding of political psychology.

The Central Assumptions of the Field

Perhaps rather surprisingly, a number of traditional approaches within political science give psychology short shrift. Many of the theories one encounters when one first studies political science tend to emphasize the importance of structures, context, or what might be called “the nature of the times,” rather than analyzing the properties of actors or individuals. Marxism, for instance, offers an especially stark example of this tendency. It tends to discount the role of individuals in history, ascribing to material factors a powerful causal effect that overwhelms the significance of particular individuals. History, according to this dialectical view, follows a familiar and predictable drumbeat no matter who the actors involved happen to be at any given time. Within international relations theory—to give another example from a wholly different theoretical tradition—the approach called neorealism argues that we can explain a great deal about how and why a state behaves as it does by looking at that nation’s position within the international system. Superpowers, neorealists argue, tend to behave the same way no matter who they are, as do all middle powers and weak powers. If this is so, it follows that we need not trouble ourselves with the analysis of who is leading a particular state or what the leader’s psychological characteristics

happen to be. Nevertheless, political psychology has always had a special appeal for those who believe that political actors—their beliefs, past life experiences, personalities, and so on—*do* make a difference. It attracts those who believe that individual actors matter; that history is not just the story of how structures and contexts shape behavior but of how individuals can themselves shape history and politics. This is perhaps the key assumption that underlies the whole field and brings its adherents together, in spite of the great diversity of approaches within it and the equally great range of topics that political psychologists address.

A second uniting assumption is the devotion of political psychologists to what has been termed *Homo Psychologicus*, as opposed to *Homo Economicus* (Houghton, 2009; Iyengar & McGuire, 1993). Once we decide that individuals *do* make a difference—in other words, that their decisions matter in the sense of having a meaningful impact on historical outcomes—we need to adopt some view of *how* they decide. Two rival models of decision making have come to dominate thinking about political behavior within modern political science, one derived from economics, the other from psychology. These are summarized below.

Homo Economicus

- Humans are comprehensively rational actors.
- Decision makers are assumed to possess perfect information.
- The decision maker generates a list of all available options.
- He or she weighs up the costs and benefits of various options.
- He or she then selects the alternative that delivers the greatest benefits relative to cost (maximizes subjective utility).
- This model is derived from microeconomics or classical economics.

Homo Psychologicus

- Humans are *boundedly rational* actors (defined below).
- Decision makers possess only imperfect information, and there are limits to everyone's cognitive processing capabilities.
- The decision maker employs various cognitive shortcuts when generating a list of available alternatives.
- Not all conceivable alternatives are fully considered.
- The decision maker selects that alternative that "will do" (in other words, the actor *satisfices* instead of maximizing utility).
- Group and broader social pressures may lead decision makers to behave in nonrational ways, even contrary to their beliefs and values.
- This model is derived from social and cognitive psychology.

Although *Homo Economicus* offers a useful set of assumptions for some political scientists—its great

strength is that it simplifies human behavior in a way that makes it predictable, and thus it appeals to those who want to model political behavior in a simplified, parsimonious way—it is not properly considered an approach to political psychology. As its name suggests, many economists and devotees of the rational choice approach to political science use it as a set of simplifying assumptions *in the full knowledge* that these assumptions do not describe how people behave in the real world; they are, however, prepared to sacrifice a measure of accuracy in the expectation that doing so will generate powerful models and predictions. However, even some economists have begun to question the utility of simplifying reality this way (a school of thought often known as *behavioral economics*). What unites devotees of a political psychological approach is precisely this reaction against oversimplification. Political psychology as a field is highly *empirical*: It is concerned with describing and explaining how political agents actually do behave, and not primarily with how they ought to, or with making simplifying assumptions for the sake of parsimony. Of course, taking this approach makes things messy; as soon as the complexity and greater realism of *Homo Psychologicus* are conceded, it becomes clear that much of human behavior is idiosyncratic and unpredictable. This is, however, a price most political psychologists are prepared to pay.

The pioneer in developing the more realistic account of human decision-making behavior called *Homo Psychologicus* here was a brilliant and eclectic academic figure known as Herbert Simon. Simon came up with at least two highly significant concepts with which he will always be associated: *bounded rationality* and *satisficing behavior* (Simon, 1957). Human decision makers are rational, he suggested, but only within the bounds of the information available to them (which is often either limited or too great to process). As a consequence, we often *satisfice* instead of maximize utility. In other words, we frequently just plump for *the first acceptable option that will do* out of a potentially limitless set of choices. So, for example, when you have not already decided where to eat one evening, you usually do not walk up and down the entire length of the street looking each place over and comparing prices and quality in minute detail; instead, you generally pick the first place that is satisfactory. And this, on a slightly different scale, is what policymakers often do, according to the bounded rationality perspective: faced with a potentially limitless range of solutions to a problem, they choose the first available option that is acceptable rather than trying to consider everything. Cognitive psychology has built considerably on Simon's early insights, and we will return to these issues when we consider the impact of that field on political psychology and the study of decision making.

What Political Psychologists Study

There are many different subfields, specialisms, and approaches within the general field of political psychology. Moreover, there are various (rather different) ways in

which an undergraduate or graduate course in political psychology may be taught. One important distinction is that some political psychologists are primarily interested in elite-level behavior. This camp focuses on examining how the perceptions of leaders shape government policies, for instance, or the impact of personality and beliefs on leadership, or on how a particular government decision came to be reached. Other political psychologists are more interested in mass-level behavior, on the other hand, or—put more simply—in how ordinary people behave. A member of this second group might study why people vote the way they do, for instance, or might be interested in the impact of public opinion on government policies or the existence of racism within a given population.

For some academics, the study of political psychology is virtually synonymous with the analysis of U.S. voting behavior, political tolerance, and the impact of the mass media on behavior. Other students of the field look mostly at foreign policy decision making and applications of psychological approaches to international relations. In truth, however, political psychology encompasses all these topics and more. One drawback of this breadth—which essentially derives from the fact that the subject matter of political psychology covers *all* varieties of political behavior—is that experts in one area of the field rarely consider themselves expert in more than one or two of the others. Nevertheless, the field of political psychology today covers topics as diverse as political communication, terrorism, genocide, the mass media, racism, emotion, cognition, neuroscience, group processes, belief systems, personality studies, and political leadership.

How the Field Is Studied

Just as political psychology encompasses an extraordinary array of topics in its subject matter, the field is equally diverse theoretically. Some of its members draw primarily on social psychological theories, for example. This large set of theoretical approaches tends to emphasize the impact of *social* situations on behavior. Other political psychologists are more influenced by cognitive psychology and the older tradition of abnormal psychology, both of which stress the importance of individual characteristics in shaping the way that we behave. Also increasingly prominent within this camp is the increasing number of political psychologists who employ the theories and methods of cognitive and social neuroscience in their work (see the section titled *Where the Field Is Going*).

In terms of the methodologies that political psychologists employ, the field has traditionally been characterized by what social scientists call *methodological pluralism*; in other words, political psychologists have used a variety of methods, both qualitative in nature (including case studies and literature reviews) and quantitative in character (most notably, large-scale survey research combined with the use of statistical procedures). Until recently, there was little evidence that any one method was predominating in the

literature, although this appears to be changing. An increasing proportion of the work published in the field's flagship journal *Political Psychology* in recent years, for instance, has been quantitative in nature, to some extent crowding out the presence of historical case studies and other qualitative work. It is unclear, however, whether this is a real trend within political psychology as a whole or whether it simply reflects an apparent preference, among recent editorial staff on the journal, for quantitative work (Monroe, Chiu, Martin, & Portman, 2009).

Origins and Historical Development of the Field

Political psychology is comparatively new as a recognized academic field. With only a few exceptions, courses in political psychology were not offered at most U.S. and European universities until the 1970s, and it was only at about the same time that the term began to be used by researchers. A *Handbook of Political Psychology*, the first of a subsequent series, appeared in the early 1970s (Knutson, 1973). A professional apparatus began to be created around the subject in the late 1970s, when the International Society for Political Psychology (ISPP) was founded. The organization remains vibrant today, and the ISPP holds its meetings as far afield as Portland, Oregon, and Barcelona, Spain. A new journal—appropriately titled *Political Psychology*—was also set up in 1979, and the field is now recognized as an integral subdiscipline within political science. While the term *political psychology* is less used within the mother discipline of psychology—the majority of adherents of political psychology continue to be employed by departments of political science—the ISPP now also includes within its ranks many professionally trained psychologists, as well as policymakers and the members of policy think tanks and nongovernmental institutions. Measured by the institutional affiliation of authors contributing to the journal *Political Psychology* since 1979, approximately 45% of all political psychologists are professional political scientists, and about 33% work in departments of psychology (Monroe et al., 2009).

The roots of political psychology run much deeper than its recent acceptance as an academic field would suggest, however. In a sense, its subject matter is as old as the study of politics itself, for as long as people have reflected on political questions, they have asked themselves basic psychological questions having to do with why human beings think and act the way they do. One of the first things one discovers in introductory political theory classes—where conventionally we consider the history of political thought as having begun with Socrates, Plato, and Aristotle—is that every political worldview is ultimately based on a view of human nature. In a general sense, every theory of politics is predicated on some general psychological portrait of how human beings are. The 16th-century Italian conservative theorist Niccolò Machiavelli, for instance, developed a famously dark view of human psychology,

which led him to propose in *The Prince* that the end justifies the means and that leaders must be prepared to do anything necessary—including committing acts of murder—to stabilize the state. Classical liberalism, on the other hand—often represented in introductory political theory courses by John Locke and Jean-Jacques Rousseau—is rather optimistic about human nature, leading to a far more benevolent idea of the role government ought to play. In the late 19th and early 20th centuries, these general conceptions of *political man* and arguments about human nature began to gradually coalesce into something more sophisticated, especially as psychology developed into a recognized academic discipline in its own right. In France in the 1800s, for example, conservative thinkers such as Hippolyte Taine and Gustave Le Bon began to develop psychological explanations of human political behavior.

The greatest contributions to the early growth of the field would come from Vienna and Frankfurt, however. Thinkers such as Sigmund Freud and Erich Fromm in particular would have a special impact on the development of the field in the United States, and Freud may in some ways be thought of as *the* founding father of the field because of his early impact on psychology and the imprint of his ideas on early work within political psychology. *Freudianism*—or *psychoanalytic theory*, as it is sometimes called—analyzes the drives or motivations that are assumed to lie within all human beings. Freud argued that sex and aggression are the most significant drives within us, but he also accorded a key role to what he called the *unconscious*—a term he virtually invented but which is now widely used in everyday speech—arguing that many of our true drives and motives are hidden even from ourselves. Because the public display of our basic drives is deemed unacceptable in many societies, their existence is often repressed. They reveal themselves only through slips of the tongue (the now famous *Freudian slips*, another term that has entered the English language) and the analysis of dreams, a medium that Freud regarded as the playground of the unconscious. He also saw the human mind as a continual battleground between our selfish, childish impulses (what he termed the *id*) and our higher, moral selves (the *super ego*). We must listen to both impulses, Freud argued. When we have a difficult time reconciling the impulses of both *id* and *superego*, however, we often subconsciously employ one or more *defense mechanisms*. They include *displacement*, *denial*, *repression*, and *transference*.

The field of political psychology has evolved through a number of fairly distinctive although overlapping historical phases during the past 80 years or so (McGuire, 1993), and we can identify three broad phases in its development: (1) the *era of personality studies* in the 1940s and 1950s, dominated by psychoanalysis; (2) the *era of political attitudes and voting behavior studies* in the 1960s and 1970s, characterized by the popularity of behaviorism and cognitive consistency theory; and (3) an era since the 1980s and

1990s, which has focused on *political beliefs*, *information processing*, and *decision making*, has used *schema theory* and *attribution theory* in particular, and has had a particular (although not exclusive) appeal for scholars of international politics. These categories will be drawn on loosely in the discussions that follow in order to show how political psychology has changed and evolved over time.

Personality Studies and Psychoanalysis

Within the United States, what would become the modern field of political psychology was pioneered during the 1920s by followers of Freud such as Charles Merriam and his student Harold Lasswell at the University of Chicago. The modern study of political psychology is generally agreed to have begun with a focus on personality studies and the appearance of several works of what is usually termed *psychobiography*, an early and still vibrant approach to studying leadership. Psychobiography focuses on the personality characteristics of political leaders and on how these characteristics affect their performance in office. Freud himself authored one or two psychobiographic works, but after his death in 1939, his primary impact on the genre came via the influence of his general theoretical approach.

Freud's emphasis on the role of unconscious motives, childhood development, and compensatory defense mechanisms would have a particular effect on the early work of Lasswell and his own student, Alexander George. It is probably fair to categorize Harold Lasswell as the first modern U.S. political psychologist because it was he who—despite initial indifference toward his ideas within the discipline—did most to probe the relationship between politics and psychology early on. Lasswell's book *Psychopathology and Politics*, published originally in 1930, now stands out as a landmark publication within the field of political psychology, as does *Power and Personality*, a now better-known work of his that first appeared in 1948. Heavily influenced by Freudian psychoanalysis, Lasswell came to argue that what he called the *political personality* results from the displacement of private problems onto public life. Simply put, Lasswell was suggesting that individuals who went into politics were often seeking political power as a compensation mechanism, seeking votes and the attention of an audience (for instance) as a replacement for love that had been lacking at home during their earlier lives.

Alexander George and Juliette George's (1964) *Woodrow Wilson and Colonel House* was similarly influenced by this kind of approach. Although not couching their analysis in especially Freudian terminology, George and George trace much of Woodrow Wilson's adult political behavior to his childhood experiences at the hands of his father, Dr. Joseph Wilson, supposedly a stern Presbyterian minister who rarely showed his son affection or congratulated him on his various achievements in life. As an adult, Wilson was propelled into

a series of conflicts with father figures of various kinds, George and George argue, and he sought the love of the people of the United States as a kind of compensation. The fame and controversy of *Woodrow Wilson and Colonel House* in turn influenced a whole host of works, such as Doris Kearns Goodwin's (1976) *Lyndon Johnson and the American Dream*, Betty Glad's (1979) *Jimmy Carter: In Search of the Great White House*, and James David Barber's (1972/1992) *The Presidential Character*, and the psychobiographic tradition remains a vibrant if (somewhat) diminished one within political psychology today.

The early impact of psychoanalysis on political psychology can also be seen in the popularity of *authoritarian personality theory* during the immediate post–World War II period. Theodor Adorno and his colleagues, who originally developed this theory, believed that right wing authoritarianism—racism and fascism, in essence—were essentially the result of rigid parental discipline within the family (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950). Authoritarian persons direct their aggression toward other groups, often racial minorities, in an attempt to compensate for a feeling of personal weakness and insecurity. The compensation mechanisms include a search for absolute answers, excessive conformity, submissiveness to authority, intolerance toward others who are unlike themselves, superstition, stereotyped thought patterns, and an oversimplistic view of reality in general (a tendency, in other words, to see things in black-and-white terms, with no shades of gray permitted). It is easy to see how the authoritarian personality approach might be used to explain the events that led up to the Holocaust, and the theory enjoyed significant popularity until the work of social psychologist Stanley Milgram suggested that we are *all* capable of extreme and unethical behaviors (Milgram, 1974). Adorno and his associates argued that obedience to authority (such as the government) will vary with one's upbringing and that rigid parental discipline had been especially prevalent in countries such as Germany during the 1920s and before.

The Rise of Voting Behavior and Research on Political Attitudes

The influence of Freud on political psychology would wane over time, however. Since it must be admitted that psychology has mostly influenced political science rather than the other way around, trends within political psychology have in general tracked changing fashions within the mother discipline of psychology. During the 1950s and 1960s—and following closely on the heels of similar trends within psychology as a whole—the influence of two other (distinctly non-Freudian) approaches would shape research within the field of political psychology: *behaviorism* and *cognitive consistency theory*. As survey techniques became more sophisticated—making it possible to ascertain the

attitudes and opinions of large numbers of people—attention would also turn from analyzing only political elites to the examination of mass political behavior.

During the postwar period, psychologists such as B. F. Skinner—a devotee of the school of psychology known as behaviorism—began to highlight what they regarded as the fundamentally unscientific nature of Freud's work. Skinner criticized Freud for focusing on untestable propositions (Skinner, 1953). Proper science, Skinner believed, ought to focus on what is testable and measurable, and what is testable and measurable is behavior (in other words, what we can see and quantify). We cannot see or measure people's thoughts, and any attempt to do so—particularly speculation of the sort Freud had engaged in—was bound to lead to bogus science, Skinner argued. At the same time, a behavioral movement within political science began to challenge a reliance on qualitative or heavily descriptive inquiry, arguing that a science of politics could be built only via the patient accumulation of data and the rigorous testing of theories against those data. Heinz Eulau's (1963) book *The Behavioral Persuasion in Politics* was emblematic of this movement, a tradition which remains strong today. Although it is hard to say in retrospect which came first or what precise impact Skinner's ideas had on political psychology—and it must be conceded that there is no “Skinnerian” movement within political psychology comparable to the one animated by Freud—his ideas were at least strikingly similar to those espoused by many students of mass behavior and (more generally) to the behavioral movement within political science. Large-scale survey research and a focus on behavior and on what is quantifiable, rather than the qualitative analysis of particular individuals, became the preferred method of the day for many political psychologists during the 1960s and 1970s, and this remains true today.

At the same time, the older psychoanalytic tradition was challenged on another front, one which—while retaining the Freudian notion of denial, or the rationalizing away of the facts as a central cognitive mechanism humans engage in—dispensed with the idea that Freud's was somehow associated with abnormal development during childhood. According to the theory of *cognitive consistency*, inconsistencies between our beliefs—or between our beliefs and our behavior—cause us to experience an uncomfortable state of tension, at least if we are made aware of our inconsistencies. Social psychologist Leon Festinger (1957) famously referred to this condition as *cognitive dissonance*, a term which has since entered the English language (though it is not always used in precisely the way he intended). Since we generally do not like to be inconsistent, we become motivated to reduce dissonance in some way and bring things back into balance or consonance.

For readers unfamiliar with this approach, the Marian Keech story may prove illuminating and, it is hoped, amusing as well (Festinger, Riecken, & Schachter, 1964). During the 1950s, Festinger infiltrated a religious cult

whose leader, “Marian Keech,” was predicting the end of the world (her name was changed in Festinger’s book to protect her identity). Keech specifically predicted that the world would come to an end on December 21, 1954, but she also prophesied that a flying saucer would pick up the “true believers” on December 20, thus saving them from all the death and devastation that would befall the rest of the world. Many members of the group had invested a lot in Keech’s prediction: They had given up their jobs, given away their savings, and sold their houses in preparation for the coming of the flying saucer. For Festinger, this represented a tantalizing scenario for the testing of his theories. He knew—or at least strongly suspected!—that the flying saucer would never show up, and this in fact proved to be the case.

But what fascinated Festinger was what the group members would do when their theory proved false. What would they do? When the saucer failed to show, Keech had a new (and rather convenient) “vision from God” shortly before 5 a.m. on the 21st, saying that “everyone was saved.” The group members then rationalized away the evidence that they had been wrong all along, and for some, the saucer’s nonappearance even strengthened their belief in the cult! While it would be easy to dismiss the members of the group as simply crazy, Festinger thought that this incident actually illustrates a very common and very human psychological tendency. While *Homo Economicus* suggests that we just update our beliefs when new information becomes available—correcting theories that have been shown to be incorrect—Festinger argued that in reality we usually just ignore or try to explain away dissonant information somehow. We bring things back into balance, in other words, by coming up with some sort of psychologically comforting excuse.

As political psychology turned from an exclusive focus on elites and began to concentrate more on mass behavior, cognitive consistency theory—an approach to psychology that, as we have suggested already, is explicitly suited to the study of attitudes and beliefs—played a central role in the most popular theory of voting behavior developed during the 1960s: the *party identification approach*. This was originally proposed by Angus Campbell and his colleagues (Campbell, Converse, Miller, & Stokes, 1960) at the University of Michigan in their book *The American Voter*. Some studies of voting in the immediate postwar period had suggested that social and economic factors directly determined our voting behavior, so that we can expect a rich man to vote Republican, a poor one to vote Democratic, and so on. But Campbell and his colleagues argued that the picture is more complex than this: A psychological variable, which they called *party identification*, plays an intervening role between “objective” social forces and the way we vote. During our formative years, Campbell and his associates proposed, we develop a long-lasting, stable attachment to a particular political party. Once formed, this loyalty becomes difficult to change and can take the form of an almost religious devotion to “our” party.

How was this approach influenced by cognitive consistency theory? Put simply, it suggested that strong partisans

simply screened out or rationalized away unfavorable information about their own party. These *strong identifiers* were so attached to their party that in some cases they would even end up voting for a party they did not agree with in an ideological sense! During the mid-1960s, for instance, the Democratic Party embraced the cause of civil rights for African Americans, a measure many Southern Democrats opposed at the time. However, substantial numbers of Southern Democrats continued to vote for the Democratic Party for many years after this (and there are probably even today some Southern Democrats who identify with the party despite an opposition to racial integration, although their numbers have certainly dwindled). Equally, many conservative Democrats continued to vote Democrat for many years after the 1930s, when the party embraced what is essentially a liberal economic agenda. Why did this occur? The work of Philip Converse (1964) in particular argued that most voters lacked an internally consistent system of attitudes and beliefs, relying instead on long-term party ties in deciding how to vote. Strong partisans would explain away their party’s poor economic performance, for instance, as the result of something other than their president’s policy choices (they might blame global economic trends, for instance). And they would ignore information about their own party’s standard bearer that did not fit the voting choice they had made.

Decision Making and International Politics

The influence of cognitive consistency theory began to be felt acutely within international relations theory as well during the mid-1970s. Robert Jervis’s (1976) best-known work in international relations, *Perception and Misperception in International Politics*, led the way in this regard, explicitly drawing on the theory of cognitive consistency to make a variety of (then path-breaking) observations about the ways in which the processing of information can fundamentally impact foreign policy decision making and outcomes on the world stage (see also Holsti, 1962). Similarly, approaches drawn from social psychology, such as the groupthink perspective of Irving Janis (1982), also had a significant impact during the same period. Janis showed how the dysfunctional processes he believed to be inherent within certain kinds of highly cohesive groups can lead to decision-making fiascos. Examining well-known episodes from U.S. foreign policy, such as Pearl Harbor, the Bay of Pigs, and the Vietnam War, Janis attributed the faulty decisions in those cases to a phenomenon he called *groupthink*, a tendency to come to a premature and ill-considered consensus within a group before all options and alternatives have been fully considered.

Indeed, during the 1970s and 1980s, cognitive approaches in general—perspectives that emphasize the content of people’s knowledge structures in shaping decision making and behavior in general—began to dominate political

psychology. This trend built on earlier work by Alexander George (1969) on the content of belief systems, a tradition called *operational code analysis*, which remains vibrant today. A whole bundle of cognitive perspectives, including *attribution theory* and *schema theory*, began to influence the field. One thing that all these perspectives share is the assumption that human beings are inherently *limited* in terms of their cognitive capabilities. Unlike, say, computers, human beings have only a limited capacity to process incoming information. We have already seen that the *Homo Economicus* model asks a great deal of human capabilities; to make a fully and comprehensively rational decision, we require *all* the relevant information pertaining to the issue we are facing. But in the real world, we know that actual human beings possess neither perfect information nor the inexhaustible energy needed to consider all alternatives. It may sound like a cliché, but the world is an incredibly complex place, and the average individual is constantly bombarded with information, not all of which can be processed efficiently or effectively.

Imagine that you want to make a fully rational, fully informed decision about where to eat tonight and that you have decided to eat out rather than at home. To meet the standard of pure rationality, you would in principle have to read all the menus of all the cafes and restaurants in your town or city. You would have to taste the various dishes in each dining option that night, comparing taste and quality and price and deciding which represented the optimal choice given your preferences. In that way, you would—as economists put it—*maximize your utility*, selecting the best option relative to its cost. Of course, in the real world, human beings very rarely behave this way. As the neuroscientist Antonio Damasio (1994) has suggested, practically the only individuals who actually make decisions in this laborious, time-consuming way are people who have experienced damage to the prefrontal cortex, an area of the brain located at just about eye level that is closely associated with emotions and decision making. In his book *Descartes' Error*, Damasio relates the story of a brain-damaged patient whom he calls Elliot. When asked to set up a time for his next appointment, Elliot begins an all-encompassing attempt lasting several hours to weigh up the pros and cons of every conceivable date in his diary until his exhausted doctors ask him to stop. As we have seen already in describing the *Homo Psychologicus* approach, what normal decision makers do instead of this is to process information by means of what are generally called cognitive shortcuts or heuristics. These are devices for prematurely cutting short the search for information, tactics that allow us to reach a reasonable decision more quickly and expeditiously than we could if we were to replicate Elliot's approach. Both schema theory and attribution theory focus on the use of such heuristics, and each has had a notable impact on the study of foreign policy decision making.

A couple of examples drawn from the literature will give you a good idea of how such heuristics work in the real

world and how they can affect both the foreign policy decisions made at the highest levels and the decisions of ordinary voters. Schema theory, for instance, argues that human beings are basically *categorizers*: Rather than considering every bit of information that comes to us afresh, we tend to fit people, events, and things into established mental “boxes” in our heads. It just so happened that when U.S. president Harry Truman and Soviet leader Joseph Stalin first met in 1946, Stalin put Truman in mind of his old boss and mentor Thomas Pendergast, a party boss from Truman's early days in Missouri (Larson, 1985). Pendergast had taught the future president the importance of keeping one's word in politics, something that had an important early influence on Truman's mind-set. Because Stalin just happened to look very much like Pendergast, Truman initially reacted warmly toward the Soviet leader and assumed that Stalin would keep his promises, just as Pendergast had. This proved to be a great error, because the Soviet leader would soon break many of the promises he had made in the aftermath of World War II. Historical analogies constitute another type of cognitive schema, and these devices have been especially well studied and analyzed within the field of foreign policy analysis (Khong, 1992).

Something rather similar to what Truman did in the Pendergast case seems to happen when voters make decisions regarding candidates they know little or nothing about. Consider what happens when we are voting for candidates in U.S. presidential primaries, for instance. We often know very little about the candidates who run for our party's presidential nomination; some may be governors of states we know little about, for instance, and even if they are members of the Senate, we may know little about them. When we are choosing between candidates of opposing parties, we can just use our party identification as a shortcut, but how do we make a decision when *all* the candidates come from our preferred party? From the perspective of schema theory, we probably just assess candidates according to how closely they fit our existing conception of the “ideal candidate.” Under such conditions, we base our voting decisions on only a few pieces of observable “data,” and we use this incomplete information to fill in what we do not know by matching a candidate to some stereotype stored in our heads (Miller, Wattenberg, & Malanchuk, 1986; Popkin, 1993). For instance, a candidate who appears “Kennedyesque”—that is, who seems to evoke the image of the late president John F. Kennedy—is likely to do quite well, whereas a candidate who evokes an image of a failed candidate is far less likely to do well at the polls.

Where the Field Is Going: A Fourth Phase?

It may perhaps be too early to talk of a fourth phase in the development of political psychology, but if recent trends are anything to go by, the field may already have entered one. This might loosely be termed the *era of emotion and*

neuroscience. This most recent trend is in large part a reaction against the computer analogy that implicitly underlies much work on cognition and decision making. For the purposes of analyzing behavior, it was often assumed, in the work described in the previous section, that human beings processed information much as a computer does, in a “cold” or neutral way. The brain was treated as little more than a storage system. This was somewhat ironic because it placed supporters of *Homo Psychologicus* and *Homo Economicus* in the same boat, in the sense that both essentially ignored the role of emotion or what some have called *hot cognition*. But human beings do not simply process information; we *feel* things as well. Virtually everything in politics—including political ideas, political issues, and politicians themselves—is loaded with emotion, either positive or negative. Very few people can look at a picture of the World Trade Center falling on September 11, 2001, for example, without feeling something, and this is a factor that obviously differentiates us from computers. Politics often provokes strong emotions in us, feelings such as happiness, sadness, anger, guilt, gratitude, disgust, joy, insecurity, fear, and anxiety.

One interesting insight that has come out of this new body of literature so far is the recognition that emotions are not necessarily irrational. For a long time, emotions have been thought of as something that comes from the heart or the gut rather than the mind. This way of thinking about our reasoning processes has been present in popular culture for hundreds if not thousands of years and probably dates back to the ancient Greeks, and it is still very common in the Western tradition of political thought to contrast ordered reason with passion or emotion. Emotion according to this view is something detrimental to informed, factually based decision making. Yet although we can all think of cases in which emotions have had a damaging impact on decision making—the fears that President Lyndon Johnson seems to have experienced during the Vietnam decision making or the feelings of depression Richard Nixon appears to have suffered during the Watergate scandal might be seen as good examples—there is an increasing recognition within political psychology that emotion is actually essential to good decision making. We have already mentioned the work of Antonio Damasio, for instance, who has shown that patients who lack the ability to feel emotion make consistently bad and even reckless life decisions. In order to reason successfully, we first have to *care* about the outcomes that might potentially result from our actions. The intense fear that we know gripped U.S. decision makers during the Cuban missile crisis, for instance, seems to have had a beneficial impact on decision making, in the sense that President John F. Kennedy eventually chose a set of options for resolving the crisis that de-escalated the conflict rather than the other way around. Within political psychology, the work of George Marcus (2002; Marcus, Neuman, & MacKuen, 2000) on political tolerance and voting behavior and the

research of Jonathan Mercer (2005) within international relations make especially prominent use of this insight about the role of emotion in decision making.

Another thing that has coincided with the recognition that humans are not much like computers is a series of technological advances in the field of neuroscience, the study of the brain. Part of the reason emotion has traditionally been neglected as a factor within decision making is that it is so hard to measure in a scientific way. Indeed, we often cannot be 100% sure what even our closest family members are thinking and feeling. Traditionally, political psychologists have relied on questionnaires and interviews to gauge what people are feeling, but these techniques are unsatisfactory in many cases, not least because people may not state honestly what they think or feel (racially prejudiced individuals may not admit to being racist in questionnaires, for instance). There has been a tendency within political psychology—a hangover, perhaps, from the behavioral era—to neglect what we cannot see or measure. Increasingly, however, we now have the capacity to “see” emotions working within the human brain. Because in recent years neuroscientists have vastly increased our knowledge of what individual parts of the brain do—we know, for instance, that the part of our brains called the *insula* is associated with disgust and the *amygdala* is activated by fear—and because brain imaging techniques such as functional magnetic resonance imaging have been developed, it is possible to directly or indirectly measure the emotions that individuals are experiencing. Political psychologists have begun to work with neuroscientists at an interdisciplinary level to use such techniques in their work. This work is very new indeed, and the results of the few studies done so far are extremely preliminary, but interesting work is already being done in this area (see, e.g., Westen, 2007). Increasing use of brain imaging techniques and an enhanced focus on the role of emotion—as well as the ways that hot processes interact with cold ones—appear to be the future of political psychology, especially in the study of mass behavior.

References and Further Readings

-
- Adorno, T., Frenkel Brunswik, E., Levinson, D., & Sanford, N. (1950). *The authoritarian personality*. New York: Harper & Row.
- Barber, J. D. (1992). *The presidential character: Predicting performance in the White House* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall. (Original work published 1972)
- Campbell, A., Converse, P., Miller, W., & Stokes, D. (1960). *The American voter*. New York: Wiley.
- Converse, P. (1964). The nature of belief systems in mass publics. In D. Apter (Ed.), *Ideology and discontent* (pp. 206–261). New York: Free Press.
- Cottam, M., Dietz Uhler, B., Masters, E., & Preston, T. (2009). *Introduction to political psychology* (2nd ed.). New York: Psychology Press.

- Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Putnam.
- Eulau, H. (1963). *The behavioral persuasion in politics*. New York: Random House.
- Festinger, L. (1957). *Theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Festinger, L., Riecken, H., & Schachter, S. (1964). *When prophecy fails: A social and psychological study of a modern group that predicted the destruction of the world*. New York: Harper & Row.
- George, A. (1969). The "operational code": A neglected approach to the study of political leaders and decision making. *International Studies Quarterly*, 13, 190-222.
- George, A., & George, J. (1964). *Woodrow Wilson and Colonel House*. New York: Dover.
- Glad, B. (1979). *Jimmy Carter: In search of the great White House*. New York: W. W. Norton.
- Goodwin, D. K. (1976). *Lyndon Johnson and the American dream*. New York: St. Martin's Press.
- Holsti, O. (1962). The belief system and national images: A case study. *Journal of Conflict Resolution*, 6, 244-252.
- Houghton, D. P. (2009). *Political psychology: Situations, individuals and cases*. New York: Routledge.
- Iyengar, S., & McGuire, W. (Eds.). (1993). *Explorations in political psychology*. Durham, NC: Duke University Press.
- Janis, I. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*. Boston: Houghton Mifflin.
- Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.
- Jost, J., & Sidanius, J. (Eds.). (2004). *Political psychology: Key readings*. New York: Psychology Press.
- Khong, Y. F. (1992). *Analogies at war: Korea, Munich, Dien Bien Phu and the Vietnam decisions*. Princeton, NJ: Princeton University Press.
- Knutson, J. (Ed.). (1973). *Handbook of political psychology*. San Francisco: Jossey Bass.
- Kressel, N. (Ed.). (1993). *Political psychology: Classic and contemporary readings*. New York: Paragon House.
- Larson, D. W. (1985). *Origins of containment: A psychological explanation*. Princeton, NJ: Princeton University Press.
- Lasswell, H. (1930). *Psychopathology and politics*. Chicago: University of Chicago Press.
- Lasswell, H. (1948). *Power and personality*. New York: W. W. Norton.
- Lau, R., & Sears, D. (1986). *Political cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- Marcus, G. (2002). *The sentimental citizen: Emotion in democratic politics*. University Park: Pennsylvania State University Press.
- Marcus, G., Neuman, R., & MacKuen, M. (2000). *Affective intelligence and political judgment*. Chicago: University of Chicago Press.
- McDermott, R. (2004). *Political psychology in international relations*. Ann Arbor: University of Michigan Press.
- McGraw, K. (2000). Contributions of the cognitive approach to political psychology. *Political Psychology*, 21, 805-827.
- McGuire, W. (1993). The political psychology relationship: Three phases of a long affair. In S. Iyengar & W. McGuire (Eds.), *Explorations in political psychology* (pp. 9-35). Durham, NC: Duke University Press.
- Mercer, J. (2005). Rationality and psychology in international politics. *International Organization*, 59, 77-106.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Miller, A., Wattenberg, M., & Malanchuk, O. (1986). Schematic assessments of presidential candidates. *American Political Science Review*, 80, 521-540.
- Monroe, K. (Ed.). (2002). *Political psychology*. Mahwah, NJ: Lawrence Erlbaum.
- Monroe, K., Chiu, W., Martin, A., & Portman, B. (2009). What is political psychology? *Perspectives on Politics*, 7, 859-882.
- Neuman, R., Marcus, G., MacKuen, M., & Crigler, A. (Eds.). (2007). *The affect effect: Dynamics of emotion in political thinking and behavior*. Chicago: University of Chicago Press.
- Popkin, S. (1993). Decision making in presidential primaries. In S. Iyengar & W. McGuire (Eds.), *Explorations in political psychology* (pp. 361-379). Durham, NC: Duke University Press.
- Redlawsk, D. (Ed.). (2006). *Feeling politics: Emotion in political information processing*. New York: Palgrave Macmillan.
- Sears, D., Huddy, L., & Jervis, R. (Eds.). (2003). *Oxford handbook of political psychology*. New York: Oxford University Press.
- Simon, H. (1957). *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*. New York: Wiley.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Vertzberger, Y. (1990). *The world in their minds: Information processing, cognition, and perception in foreign policy decisionmaking*. Stanford, CA: Stanford University Press.
- Westen, D. (2007). *The political brain: The role of emotion in deciding the fate of the nation*. New York: Public Affairs.

8

STRAUSSIANS

SCOT J. ZENTNER

California State University, San Bernardino

Leo Strauss was one of the most prominent and controversial political theorists of the 20th century. He is perhaps most well-known for his view that classical political science, exemplified by Plato and Aristotle, is superior to modern political science in its various forms. Strauss cultivated in his students and admirers a certain disdain for contemporary political science, which he believed was largely irrelevant or even dangerous to political life. He emphasized the need for political science to be prescriptive with respect to the ends as well as the means of political action.

Strauss's followers are now commonly known as the *Straussians*, although some of them resist the label. While there are disagreements among them, they generally adhere to his rejection of mainstream political science, with its emphasis on method, math, and theory. They make up a relatively small but important group within academic political science, several holding posts in some of the most prestigious universities in the United States. While most of them hold formal positions in the field of political philosophy, their work extends to all the substantive fields of contemporary academic political science.

In what follows, the Straussian approach to political science and the place of the Straussians within the discipline are examined. First, the Straussian preference for classical political science is explained. Second, the array of Straussian scholarship is reviewed, particularly with a view to its practical ends. Third, given the political emphasis of the Straussians, the political reaction to their work is

explored. Finally, the political and philosophic divisions among the Straussians are examined.

Classical and Modern Political Science

The Straussian approach to political science may be understood in light of what Strauss viewed as the crisis of liberal democracy. The crisis consists in the loss of confidence in the principles underlying liberal democracy (Strauss, 1964). This loss of confidence was largely caused, Strauss argued, by tendencies manifested in the practical application of those principles, including an excessive preference for rights over duties, a naive and even dangerous belief in progress, and an inclination to moral relativism. Stated differently, liberal democracy had become too liberal and too democratic. Contemporary academic political science, Strauss argued, merely reflects this more general crisis. Strauss and his students believe that classical political philosophy and science may provide a needed corrective to these problems.

The lengthiest Straussian critique of academic political science is *Essays on the Scientific Study of Politics* (Storing, 1962), a collection of articles written by Strauss and several of his students. The essays are sharply critical reviews of several major works in *behavioral political science*, which the authors term the *new political science*. The book provoked a vigorous response within the discipline. John Schaar and Sheldon Wolin (1963) wrote a lengthy,

well-known critique. They were taken aback by what they took to be the “unrelievedly hostile and destructive” temper of the book (p. 126). Even Strauss’s supporters have noted the “bellicose” style of his memorable epilogue to the volume (Behnegar, 2003, p. 143). The book as a whole, and Strauss’s essay in particular, may be taken as the quasi-official declaration of the distinctively Straussian approach to political science.

The Critique of Science

In the epilogue to *Essays*, Strauss criticizes behavioral political science from the classical point of view, which he locates in the thought of Aristotle. First, the new political science drops the Aristotelian distinction between theoretical and practical sciences. The practical success of modern physics and mathematics, which in themselves are theoretical, led to the adoption of those sciences as the models for the practical sciences, including political science. In this way, the sciences as such came to be thought of as distinct from and superior to philosophy. Second, the blurring of the practical and theoretical sciences leads to the distortion of political things, which properly concern matters of action rather than contemplation. The classical view, moreover, held that political action contained principles of its own, irrespective of the claims of theory. The political sphere may require a theoretical defense from the intrusions of visionary or theoretical claims, but politics properly understood is altogether practical in nature (Strauss, 1962).

Third, since politics falls into the realm of practical science, political scientists cannot be mere neutral observers of political action. Political scientists, according to the classical view, must be directly engaged in political affairs. But since political scientists are more learned in the nature of political things, they become “umpires” of political disputes; they are more than mere partisans in those disputes. The new and more theoretical political science, on the other hand, seeks to refrain from such activity as a matter of scientific neutrality. Fourth, because of their political role, classical political scientists necessarily evaluate political phenomena. The new political science, in contrast, precisely insists on not making value judgments. Its scientific neutrality assumes the separation of facts from values; only the former are the proper objects of scientific analysis.

The source of the stridency of the Straussian critique of the new political science may be found in Strauss’s (1962) summary claim that “the political is *sui generis* and cannot be understood as derivative from the sub-political” (p. 311). The new political science deprives human beings of their dignity, according to Strauss, because it does not understand those beings on their own terms. It understands political things as the consequence of subrational or sub-political forces or causes. The political and subpolitical realms are blurred. Classical political science, in contrast, assumes that human beings occupy a distinct place, that is, the political community, within a greater whole. As a

discrete part within the greater whole that is the world, the political community comes to sight as a whole on its own. It is this wholeness of the political community, in turn, that is the ground of the common good. Human beings as citizens in the community share something in common with one another that is different from what those same human beings share with the larger animal and material world. What is shared is their common ends or purposes as political beings, beings concerned with the good, the just, the advantageous, and so on.

The new political science, on the other hand, denies the possibility of human dignity and the common good because it understands the human in the same way that modern science understands the nonhuman; it understands human beings in terms of their origins or causes rather than their ends. According to the new political science, Strauss argues, there is only a difference in degree, and not in kind, between man and beast. The new political science rejects the commonsense or prescientific understanding of the whole made up of political phenomena. This commonsense view, Strauss insists, is the necessary basis of any practical science.

The Problem of Common Sense

Strauss (1962) argues that the obscuring of common sense renders the new political science incapable of either adding anything important to ordinary political understanding or distinguishing the politically relevant from the politically irrelevant. He argues for a genuinely empirical political science against what he considers the spurious empiricism of the new approach. But Schaar and Wolin (1963) quite reasonably point out that, although political science should be addressed to what is practically meaningful rather than what is practically meaningless, it is not always evident that common sense enables us to make that distinction. Indeed, they insist that Strauss does not provide a clear account of common sense as such. Strauss (1962) mentions, for example, the older commonsense view that assumed that witches are real. He argues that the “errors” regarding such things were eventually discovered without the aid of empiricism (p. 317). This means that on the basis of common sense, it came to be understood that people should not be accused of being witches, even though the belief in witches was originally a part of common sense. Schaar and Wolin (1963) suggest that Strauss does not really explain how this was done or what kind of knowledge common sense without belief in witches really is. They question whether, on Straussian terms, one can distinguish the truths from the errors of common sense.

The problem of common sense is central to understanding the Straussian approach to political science. Strauss (1964) remarks that his return to classical political science is “tentative or experimental” (p. 11). This tentativeness is necessary because the conditions of the ancient polity that shaped classical political science are very different from those of today. Strauss engaged in close studies of political

philosophers from the tradition in order to gain some understanding of the movement from those origins to our contemporary circumstances. This return through the reading of old books entails, among other things, awareness of the original commonsense view of witches, ghosts, and the like (Strauss, 1953). Strauss does not, of course, argue that political scientists today actually should believe in witches. However, he clearly suggests that common sense is not always so common. Moreover, what is common in the modern world is an ordinary life decisively shaped by science; it is a world without witches. This naturally raises the question whether Strauss's return to classical political philosophy should ever be more than tentative. Critics of the Straussians suggest that such a return is misguided, given the very different conditions of political science today.

These difficulties are compounded by the problematic character of classical political philosophy even in the conditions of the ancient world. Strauss (1964) notes that Plato and Aristotle understood politics from a philosophic point of view; they did not see the political community as it sees itself. They understood the political world according to the philosophic idea of natural justice rather than the commonsense view of religious authority. Strauss suggests that the historian Thucydides did not make this mistake and so was closer to the original commonsense view of politics than were the philosophers. Strauss appeals to Aristotle's view of common sense, yet he criticizes that view. He seems to admit that classical political philosophy appears to obscure common sense in a manner similar to modern political science.

Nevertheless, Strauss favors a return to classical political science because he believes that the classical approach, whatever its shortcomings, is superior to the modern alternatives. The classical approach is better, not because it is perfect, but because it takes seriously the questions of justice and morality that are central to political life. Modern political science, on the other hand, generally considers moral claims to be relative or subjective value judgments. This is not to say that Strauss thinks that modern political scientists do not take their own political views seriously. Indeed, Strauss contends that political beings properly so called do not consider their political claims to be merely relative or subjective; they consider their claims to be true, or they consider it necessary to present them as true. However, modern political scientists *qua* scientists exclude these political considerations. While classical political science may be unable to provide a complete or comprehensive account of the commonsense meaning of political life, Strauss suggests, modern political science ignores this meaning altogether as a matter of methodological principle.

The Analysis of Regimes

According to the Straussians, value-free modern political science generally fails to perform the central task of genuine political science: the practical study of regimes

(Strauss, 1962). Regimes are the formalized claims to rule based on different opinions regarding justice and morality. These claims are traditionally embodied in the various groups or parties (democrats, oligarchs, etc.) that arise in political communities. But these claims are almost always defective or incomplete in some way. In order to act as umpire within the community, the political scientist must articulate, at least in outline, what a complete regime without defects would be. The philosophers call this the *best regime* (Strauss, 1953). Political scientists presumably are able to articulate the best regime because of their deeper study of politics. Stated differently, Straussians aim to politicize political science as a science. This means more than using what is learned from political science for political ends. It means that political science should aim to discover those ends. In this sense, political science is really only an extension of politics.

However, political scientists do not necessarily attempt to reform the political community in order to make it identical to the best regime. Political scientists are not to act as partisans, convinced of the absolute superiority of their view. The best regime normally provides only a general measure of the health of actual political communities. It is most often useful because it shows why perfection is not possible in actual political life. For example, in Plato's *Republic*, Socrates suggests that the best regime includes a significant communist element requiring the abolition of families and private property. Moreover, this best regime requires the rule of philosopher kings who are not bound by the rule of law. Because of the extreme character of this best regime, Straussians usually interpret it to be a practical impossibility (Strauss, 1953). The political scientist learns from the articulation of the best regime what the practical limits of political life are.

The best regime is the regime that is best everywhere and always. It is not bound to or limited by any transitory conditions. This means that the best regime is the regime that accords with nature or is what is right by nature. The attempt to recover the classical idea of *natural right* is one of the most important features of Straussian political science. However, because the best regime by nature is rarely practical or possible, classical political science articulates a prudent moderation of its aims. As a practical matter, the rule of law and the tutelage of gentlemen is to be preferred to communism and the rule of philosophers (Strauss, 1953). Stated differently, natural right is diluted by or for the sake of political right. The political philosophers may influence politics, but they are to do so in a circumspect and practical manner.

Political philosophers nevertheless always stand in some tension with their political communities because they necessarily question the prevailing regime claims in light of the best regime (Strauss, 1959). The most famous example of this tension was the conflict between Socrates and Athens. Socrates claimed to be helping the Athenians, but they perceived his questioning of their beliefs to be a

threat to what they held most dear. They eventually put him to death for what they perceived to be his crimes. One goal of political scientists, according to the tradition that followed Socrates, is to improve political life without suffering his fate (Strauss, 1959). This requires that the political scientist support the political community while nevertheless criticizing it. For this reason, Straussians generally describe themselves as friendly critics of liberal democracy. While they often use classical sources in order to criticize modern politics, they defend this criticism as helpful to modern politics.

The Straussian emphasis on the practical character of political science resembles in certain respects two broad movements in academic political science that gained prominence in the discipline after Strauss's death: *postbehavioralism* and *rational choice*. Both of these approaches stress the idea that political science should prescribe political action, not merely describe and explain it. However, James Ceaser (1992) has expressed most fully the Straussian view that these approaches share behavioralism's main flaw, namely, indifference to the regime question. While these approaches may provide interesting or useful information about how to achieve the ends of political life, they do not seek to identify those ends in light of the best regime. Straussians believe political science proceeds directly from the commonsense arguments inherent in political life regarding the right and just ordering of society. The techniques of modern political science can be, at best, only instrumental to the precisely political role of political science.

Straussian Studies

Many of Strauss's students and admirers have followed his example of studying the great books of the philosophic tradition. Such studies are often viewed as propaedeutic, a necessary part of understanding the basic questions of political life. Straussians are known in particular for their very close readings of classic texts. This includes what some critics consider the controversial attempt to understand the great philosophers as they understood themselves, without modern presuppositions affecting the interpretation of the classics. In keeping with this approach, the Straussians have produced quite literal translations of many classic texts. Such translations are considered a part of both effective teaching and sound scholarship. The need to return to the classic tradition has even led some Straussians, most notably Seth Benardete (1989, 2000), to concentrate their work almost entirely on the study of ancient Greek philosophy and poetry.

Most Straussians, however, do not limit their scholarly work to the study of classical texts. They have produced work in all the substantive fields within academic political science. Despite the variety of their work, they may be said to share Strauss's view that political science should proceed

from the ordinary questions of political life and have as its aim the improvement of the political community. This political disposition is evident in the variety of Straussian studies. Thomas Pangle (2006) and John Murley (2005) have provided extensive reviews of these works. It is helpful here to highlight briefly several characteristic examples of Straussian scholarship.

Allan Bloom is perhaps the most well-known of Strauss's students. He earned popular acclaim with his best-selling *The Closing of the American Mind* (1987). The book is largely an indictment of higher education and popular culture in the United States from the Straussian point of view, particularly the idea that the classical emphasis on moral and intellectual virtue has been vitiated in the name of the modern turn to egalitarian and democratic principles. The book was especially popular for its often amusing take on the manner in which these modern tendencies manifest themselves in the social lives of college students. In this respect, the reaction to the book provides some insight into the popularity of Straussian ideas among many college students exposed to them. More broadly, the attention paid to Bloom's book reflects the attractiveness of the Straussian critique of modernity. There are a number of postmodern criticisms of modern life, but with their appeal to the philosophic tradition and common sense, Straussians offer a distinct alternative, however controversial, within the academy.

Consistent with Strauss's emphasis on common sense and traditional thought, Straussians place great importance on the study of the origins of political communities. Given that most Straussians live in the United States, this naturally has resulted in an interest in the American regime. Martin Diamond (1992) and Herbert Storing (1995), for example, emphasized the role that political ideas and statesmanship played in the founding of the United States. This approach rejects theories that find in economic or psychological forces the chief causes of the founders' actions. Straussians examine the regime claims of the various partisans of the time, and the arguments made in support of those claims, with a view to an articulation of the genuine meaning of the Constitution and other major documents of the founding.

The Constitution's influence has been a particularly important part of the Straussian study of the United States. Harvey Mansfield (1991), for example, has suggested that the forms of the Constitution serve as the principal brake on the democratic tendencies of the regime. Those tendencies are derived from the revolutionary and egalitarian origins of the United States. The maintenance of constitutional forms is, Mansfield argues, crucial to the success of the regime. In this respect, like many Straussians, he follows Alexis de Tocqueville in his study of the customs and mores that sustain or harm regimes.

Indeed, according to many Straussians, Tocqueville's work is the model for classical political science in the modern world (Ceaser, 1992). In keeping with the idea that

one should be a friendly critic of liberal democracy, Tocqueville criticized the United States and France with the intention of improving liberal democracy as the best practical form of government under modern conditions. In similar fashion, while many Straussians are critical of the underlying principles of the United States as unduly modern, they often stress what they take to be the more salutary elements of the regime, such as the adherence to constitutional forms.

Straussians generally emphasize the need for moderation in liberal democracy, particularly in light of the rise of progressive ideas in the past century. They look to the classical descriptions of the gentleman and the statesman as models for politicians today. Strauss suggested that the classical philosophic notion of aristocracy, however inapplicable to modern conditions, remained in principle the generally superior regime type. In keeping with this idea, Straussians generally argue that the aristocratic elements of liberal democracies should be encouraged if those democracies are not to suffer the traditional failings of popular governments. Supporting the aristocratic elements requires, in addition to the defense of constitutional forms, support for the cultivation of principled statesmanship. Respect for forms and the teaching of statesmanship are related aspects of regime analysis, for the holders of duly constituted public offices are the practical embodiment of the regime.

Accordingly, with respect to the courts, Straussians such as Walter Berns (1987) tend to share with other conservative scholars a respect for constitutional forms and, in particular, a preference for judicial restraint. They generally oppose the idea that the meaning of the Constitution evolves or changes over time. This opposition follows from the Straussian concern to maintain the forms of the Constitution as the moderating force in the democracy. When deciding cases, judges should look more often than they do to the wisdom of the framers of the Constitution and the great jurists of the founding era.

Straussians have examined the American presidency in a similar fashion. Jeffrey Tulis (1987), for example, has criticized the rise of the modern progressive presidency, with its emphasis on popular rhetoric and dependence on democratic opinion. James Ceaser (1979) has examined the extent to which changes in electoral arrangements have contributed to this development. These studies argue that under current circumstances, the presidency is much less able to fulfill the moderating and quasi-aristocratic role in the regime intended for it by the founders. Classical regime analysis in this case reveals the particular problem of democratic decay in the conditions of the modern state.

Other Straussians, especially Diamond and Storing, as well as their students, have taken an alternate view, but with the same intention, by locating potentially moderating institutions and practices within contemporary U.S. government. In general, they are less critical of liberal democracy in general than are most other Straussians. For

example, Joseph Bessette (1994) has argued that the modern Congress, despite its shortcomings, in many ways fulfills the founders' original hope for a deliberative legislative process. John Rohr (1986) has considered the possibility that the modern civil service, despite its significant role in the modern state, might, with the proper guidance, play a moderating or even statesmanlike role in the mature American regime.

The most significant Straussian scholar of American politics is Harry Jaffa. Among Straussian works, his writings on Abraham Lincoln, slavery, and the Civil War provide the most searching account of the relationship between political philosophy and statesmanship. In his early writings, Jaffa (1959) argued that Lincoln corrected the regime of the founding fathers, which unduly relied on the modern idea that the preservation of private rights is more important than moral obligation. In his later writings, Jaffa (2000) has presented a more subtle and complex interpretation, arguing that Lincoln did not correct the founders' modern principles but actually recovered the classical core of their thought. Jaffa insists on the need to understand the thought of the founders and Lincoln on its own terms and not simply as a by-product of the general currents of modern thought.

Straussians highlight the degree to which the ideas and principles of rulers, as opposed to economic or psychological causes, shape the course of events. As exemplified in Jaffa's work, Straussians often focus their studies on the writings and speeches, the reasoned arguments, of political leaders. This follows from the idea that political things must be taken on their own terms if we are to see them fully for what they are. Straussians have taken this approach to other areas as well, including political psychology and political economy. These studies aim to show that the phenomena in question are decisively political in character and not reducible to other causes. This is especially evident in Straussian studies of the problem of tyranny. Scholars such as Charles Fairbanks (1993) and Myron Rush (1974), for example, have followed Strauss's (1991) example by showing that modern communist and fascist tyrants are best understood in terms of their ideological self-understanding.

Also noteworthy are Straussian works in the area of international relations. In keeping with the return to classical sources, Clifford Orwin (1994) and others have engaged in close and detailed studies of Thucydides and other great books in the tradition. They have followed Strauss's lead in reading Thucydides' account of justice between nations as a companion to the works of the great philosophers. These new interpretations diverge from more traditional "realist" accounts of the Greek historian. Added to this work have been further Straussian studies of Machiavelli, Kant, and other sources of modern realism, neorealism, and idealism. These philosophical and historical readings have led Straussians to question a number of the assumptions of these more contemporary theories.

Not surprisingly, Straussians emphasize the important relationship between political regimes and the conduct of international affairs. Classical political philosophers wrote more extensively on domestic matters than on foreign affairs. From the classical perspective, policy concerns regarding war and diplomacy cannot be understood apart from regime analysis. The most famous classical example of this is Socrates' depiction of the guardian class in Plato's *Republic*. In this light, Straussians have explored the relationship between foreign affairs and the character of modern liberal democracies. Jeremy Rabkin (2005), for example, has argued that sovereignty is an essential element of constitutional democracies. Only sovereign governments can make and enforce laws needed for the protection of individual rights and the maintenance of democratic norms. Straussians generally are wary of the contemporary trend toward multilateral agreements and the like, which they believe can pose serious threats to sovereignty and constitutional government.

Politics and Political Science

Any discussion of the Straussians must address the political controversies surrounding their work. Straussians have come under criticism, some of it quite severe, for appearing to suggest that they have special knowledge of the best regime. This criticism is intensified by the political conservatism of most Straussians. A number of Straussians are or have been active in conservative media outlets and public policy institutes, as well as in the Republican Party. While not all Straussians are politically conservative, relatively few of them are liberal.

The political criticism of the Straussians became most acute during the administration of President George W. Bush, which employed a number of foreign policy assistants with varying connections to Strauss or the Straussians. The Straussians were criticized, in particular, for their association with neoconservatives, who generally have favored an interventionist and unilateral foreign policy. The harshest critics accuse the Straussians of being a cabal that deceives the public in order to justify war and undermine democracy. This charge of deception was made most often with respect to the Bush administration's claims regarding weapons of mass destruction in Saddam Hussein's Iraq.

While many of these critics may be dismissed as narrowly partisan, the controversy did highlight an important element of the Straussian approach to political science. Several academic critics of Strauss, most notably Shadia Drury (2005), have argued that Strauss believed philosophers like himself were fundamentally superior to ordinary human beings. Strauss's best regime of natural right, Drury suggests, really amounts to only a defense of the rule of the strong over the weak. The Straussians, she claims, believe they are a kind of philosophic aristocracy. They believe the

ignorant many exist only to serve the philosophic few. In this reading, Strauss follows Machiavelli and Nietzsche by glorifying fraud and power.

The principal source for Drury's charge is Strauss's well-known view that political philosophers often write in a manner that both reveals and conceals what they really think (Strauss, 1952). The philosophers do this as a matter of prudence. On one hand, the philosopher may be persecuted, as in the case of Socrates, because the people perceive the philosopher's questioning to be a threat to the community. On the other hand, the philosopher's questioning may very well undermine the necessary beliefs of the community. In this sense, Athens may actually have been justified in condemning Socrates. The solution to this problem is represented most famously by Plato, who communicates poetically through the characters in his dialogues and not directly in his own name. Only the philosophic few presumably can decipher the secret meaning of Plato's texts. In this way, the philosopher writes exoterically for the many and esoterically for the few.

Most Straussians defend the general notion of esoteric speech and writing as a part of political life itself. It is commonly understood that statesmen often refrain from fully stating their opinions. For example, as Jaffa (1959) has argued, Abraham Lincoln's rhetorical efforts to navigate between the competing forces of slavery and abolition were akin to exoteric writing. Even though he favored abolition, Lincoln could not state explicitly his real aim to end slavery in the South. Similarly, given the racist opinions of his time, he could not openly support social and political equality for African Americans. Lincoln's calculated rhetoric, Jaffa argues, was intended to serve practically the ends of equality and freedom rightly understood. Straussians similarly defend their political science as a form of prudence.

Political scientists of almost every stripe, including Drury, admit that some degree of elitism, if only for the sake of competence, is necessary and even desirable in liberal democratic politics. The larger question concerns the appropriate degree and form of elitism. At various places in his writings, Strauss refers to a fundamental conflict between philosophy and politics. Because of their greater knowledge, philosophers tend to form a class separate from the rest of the community. Because of this, politics and philosophy cannot be completely reconciled. Moreover, the separation of politics and philosophy in some sense is necessary for the proper understanding of both.

The Straussian critique of modern political philosophy centers on the modern attempt to overcome the separation between politics and philosophy. Classical political philosophy proceeded from the view that the many could never become truly philosophic. Modern philosophy, on the other hand, attempted to make the many philosophic, in a sense. It did this by means of enlightenment. But in order to enlighten, modern philosophers had to simplify political life. Human things had to be portrayed in nonhuman terms.

This is, as mentioned above, the root of the Straussian critique of contemporary political science. Strauss argued that the simplification of human things brought about a distortion of political philosophy. A characteristic trait of genuine political philosophy is awareness of the immense complexity of human things. Political philosophy lacks a complete account of all things and so cannot rule political life in a comprehensive way. Stated in Straussian terms, natural right cannot completely replace political right. Strauss argues that modern philosophy's simplification of human things through enlightenment is, in effect, an attempt at such replacement. It is the real effort to subordinate politics to philosophy.

A number of Straussians have defended Strauss from the charges by Drury and others. Catherine and Michael Zuckert (2006) have presented the lengthiest such defense. The novelty of their argument is the claim that Strauss did not, in fact, write esoterically. In particular, they claim that Strauss did not have a secret teaching. Strauss is known for often writing in a cryptic and enigmatic manner. This has naturally led many observers to think, in light of his explicit descriptions of esoteric methods, that Strauss also wrote esoterically. The Zuckerts, however, argue that Strauss's writing style indicates his "pedagogical reserve," not his esotericism (p. 136). Modern liberal democracy, with its relative degree of openness, has largely eliminated the need for esoteric writing. But while esoteric writing is no longer necessary, the need for restraint in philosophic education of students remains. Complete openness at the outset of such education might undermine students' emerging appreciation of the peculiar pleasures of philosophy. Strauss's manner of writing, the Zuckerts suggest, does not serve a political purpose.

Many observers, including many or most Straussians, remain unconvinced by the Zuckerts. Given the fact that any esoteric teaching would necessarily be difficult to discern, it is difficult to conclude that Strauss did not have one. In any case, it is not clear that such a teaching, even if it existed, must be hostile to liberal democracy. While it is fair to say that much of their work is critical of democracy in its contemporary form, it is also reasonable to conclude that the Straussians do not intend to overthrow democracy.

Straussians, East and West

The Straussians at times appear to disagree among themselves as much as they do with the wider academic and political communities. The differences among them illustrate several key aspects of the Straussian approach, particularly their understanding of the relationship between liberal democracy and political science. The Zuckerts have noted at least three distinct types of Straussians, typically identified according to the geographical locations of their principal representatives. The "East Coast" Straussians, who appear to be the greatest in number, are the most critical of liberal

democracy, which they see as too modern in its origins and principles. Prominent East Coast Straussians include Bloom, Berns, and Pangle. Jaffa and his admirers are the "West Coast" Straussians, who vigorously defend the regime of the American Founding and deny that its principles are fundamentally modern. A third group, which the Zuckerts call the "Midwest" Straussians, believes that the American Founding is fully modern but that modern principles are not as bad as other Straussians tend to think. Diamond and Storing are the most notable members of this group. Insofar as they are the least critical of modern political philosophy, the Zuckerts note, they are the least Straussian of the three groups.

The sharpest exchanges between the Straussians have been initiated by Jaffa (Zentner, 2003). He has criticized Diamond, Berns, Pangle, and Bloom, as well as prominent conservative admirers of Strauss such as Irving Kristol and Wilmoore Kendall. He insists that both sound public policy and genuine political philosophy require adherence to the principles of the Declaration of Independence. Most Straussians, he argues, wrongly apply to the American Founders Strauss's critique of early modern philosophers such as John Locke. Strauss criticized Locke for favoring rights over duties and self-preservation over virtue. Jaffa suggests that while Strauss's account of Locke is perhaps at the deepest level true, it does not apply to the founders since they did not read Locke in the manner that Strauss did. Their thought, Jaffa insists, must be understood on its own terms.

Jaffa's Straussian opponents argue that he wrongly characterizes their critique of the American Founders as more extreme than it really is. They follow, in their view, Strauss's example of the loyal critic of liberal democracy. Moreover, they suggest that Jaffa's approach lacks the skepticism required of the political philosopher. Berns, for example, suggests that Jaffa assumes political philosophy should be edifying. According to Berns, Strauss specifically denied this role to philosophy. Pangle has similarly characterized Jaffa as much too strident. The political role of the philosopher, he argues, should be more circumspect. Political philosophers necessarily follow Socrates' example and question the orthodox opinions of their fellow citizens. Political philosophers in the U.S., then, must call into question, however gently, the founders' egalitarian and democratic principles.

The quarrel among the Straussians derives, in part, from their opposed interpretations of the core of Strauss's thought. Strauss (1997) considered the "theological-political problem" the theme of his life's work (p. 453). This problem, briefly stated, concerns the tension between particularistic identities, on one hand, and universalistic commitments, on the other. Strauss observed that this was a perennial human problem. Every human being, in both ancient and modern times, is at once a natural being and a citizen, a universal being and a political being. The tension between these two aspects of life became acute with the rise of biblical religion and its universal moral beliefs.

Before biblical religion, there rarely was an explicit division between political and religious authority. All authorities in the ancient city were political or particularistic. The ancient city was what we today call a closed society. In that world, the awareness of the human or the universal in a meaningful sense became evident, if at all, only when philosophers appeared and began to articulate the universal idea of nature. Biblical religion, especially Christianity, was unlike the other ancient religions because it exhibited a universalism similar to philosophy. In this way, the relationship between politics and religion became confused or incoherent. Religion no longer provided the authority for politics in an unproblematic way.

Indeed, the West became the arena for the working out of the two universalistic claimants to temporal and political authority: Greek philosophy and biblical religion. Strauss (1997) suggested that this tension between reason and revelation was the vitality of the Western world. The tension largely emerged out of the philosophic attempt to provide or restore theological-political coherence to human life. Sometimes, as in the case of medieval natural law doctrines, this philosophic attempt actually appeared in the guise of religion. Strauss was somewhat more sympathetic, however, with modern philosophers, such as Spinoza, who attempted to provide through liberal democracy something like a commonsense ground for political life apart from religion.

Liberal democracy was intended to solve the theological-political problem by separating the state from society and the church from the state. Religious beliefs as well as different forms of prejudice, racial and otherwise, would be left to the private realm. The public realm would be dedicated to the shared concern for the protection of individual rights and a limited public good. But Strauss (1997) emphasized two basic problems with this solution. First, prejudice could fester and become more dangerous precisely because it was left alone in the private realm. Second, the separation of church and state leaves the state without the support of the highest and most binding authorities: God and divine law. The result would likely be an anemic state that could not hold sway over a society moved by prejudices of one sort or another. Strauss interpreted the collapse of the Weimar Republic in Germany and the rise of the Nazis in just this way. In related fashion, he likened the predicament of African Americans in the United States to that of the Jews in Germany.

The disagreements among the Straussians regarding liberal democracy and the American Founding reflect Strauss's ambivalence about both. The East Coast Straussians appear to have adopted the more sober aspect of Strauss's view. Although they support liberal democracy in general, and the U.S. constitutional system in particular, they tend to emphasize the need to correct what they take to be their defects. This approach is in line with the more general Straussian antipathy toward modernity. The Midwest Straussians, it should be noted, do not differ

substantially from the East Coast Straussians with respect to their general interpretation of liberal democracy and the American Founding. However, they evaluate both more positively. The emphasis on private rights and self-preservation has, they suggest, resulted in a more or less decent system of government broadly consistent with the human good. Liberal democracy appears to be particularly favorable in comparison with other available forms of government in the modern world.

The West Coast Straussian view represents the clearer innovation on Strauss's theme. Jaffa's interpretation of the American Founders and Lincoln seems tailored to address the particular problem of liberal democracy that Strauss describes. That problem rests on the hollowness of liberal democracy itself, particularly the public indifference to private opinion. Jaffa's interpretation of Lincoln and slavery suggests that liberal democracy need not be understood in narrowly Lockean or self-interested terms. He defends Lincoln's argument that the Founders' principles provide clear moral guidance on the question of slavery. The protections for slavery in the South were compromises necessary to secure the Union. But as compromises, they reflect the underlying moral principle of natural human equality. Jaffa suggests that the founders' principles reflect in important ways traditional elements of natural right that ultimately find their roots in the thought of Plato and Aristotle.

Many East Coast and Midwest Straussians believe that Jaffa has mixed classical and modern principles in untenable ways. For example, they often point out that classical natural right, which emphasizes duties over rights, is essentially inconsistent with modern natural right, which emphasizes rights over duties. Jaffa's response is that the emphasis on natural rights as an answer to the theological-political problem, that is, the basis for the separation of church and state, should be understood primarily in practical and not theoretical terms. In other words, he argues that he returns to the commonsense understanding of political life that Strauss urged on political scientists. The difficulty, as Jaffa admits, is that the commonsense understanding of practical life today is greatly influenced by theoretical considerations. This is nowhere more evident than in the fact that the idea of natural rights, on which the founders base their politics, is a product of theory and philosophy. The East Coast and Midwest Straussians emphasize this important point.

Conclusion

The Straussians are likely to continue to quarrel among themselves, not only about the principles and practices of liberal democracy, but about the meaning of Strauss's work as well. They also likely will continue to provide provocative accounts of contemporary politics and public policy, which will in turn draw further criticism. Their peculiar form of traditional political science represents a distinct, if

at times controversial, voice in the world of academic political science. However, in this the Straussians are not unusual. Their contentious views may best be seen as part of the ongoing discussions, well-known within political science, about the very meaning and purpose of the discipline.

Note: Research for this article was supported by the Social Philosophy and Policy Center, Bowling Green State University.

References and Further Readings

- Behnegar, N. (2003). *Leo Strauss, Max Weber, and the scientific study of politics*. Chicago: University of Chicago Press.
- Benardete, S. (1989). *Socrates' second sailing: On Plato's Republic*. Chicago: University of Chicago Press.
- Benardete, S. (2000). *The argument of the action: Essays on Greek poetry and philosophy*. Chicago: University of Chicago Press.
- Berns, W. (1987). *Taking the Constitution seriously*. New York: Simon & Schuster.
- Bessette, J. M. (1994). *The mild voice of reason: Deliberative democracy and American national government*. Chicago: University of Chicago Press.
- Bloom, A. (1987). *The closing of the American mind*. New York: Simon & Schuster.
- Ceaser, J. W. (1979). *Presidential selection: Theory and development*. Princeton, NJ: Princeton University Press.
- Ceaser, J. W. (1992). *Liberal democracy and political science*. Baltimore: Johns Hopkins University Press.
- Diamond, M. (1992). *As far as republican principles will admit*. Washington, DC: AEI Press.
- Drury, S. B. (2005). *The political ideas of Leo Strauss* (Updated ed.). New York: Palgrave Macmillan.
- Fairbanks, C. (1993). The nature of the beast. *National Interest*, 31, 46–56.
- Jaffa, H. V. (1959). *Crisis of the house divided*. Garden City, NY: Doubleday.
- Jaffa, H. V. (2000). *A new birth of freedom*. Lanham, MD: Rowman & Littlefield.
- Mansfield, H. C. (1991). *America's constitutional soul*. Baltimore: Johns Hopkins University Press.
- Murley, J. A. (Ed.). (2005). *Leo Strauss and his legacy: A bibliography*. Lanham, MD: Lexington Books.
- Orwin, C. (1994). *The humanity of Thucydides*. Princeton, NJ: Princeton University Press.
- Pangle, T. L. (2006). *Leo Strauss: An introduction to his thought and intellectual legacy*. Baltimore: Johns Hopkins University Press.
- Rabkin, J. (2005). *Law without nations? Why constitutional government requires sovereign states*. Princeton, NJ: Princeton University Press.
- Rohr, J. A. (1986). *To run a constitution: The legitimacy of the administrative state*. Lawrence: University of Kansas Press.
- Rush, M. (1974). *How communist states change their rulers*. Ithaca, NY: Cornell University Press.
- Schaar, J. H., & Wolin, S. S. (1963). Essays on the scientific study of politics: A critique. *American Political Science Review*, 57, 125–150.
- Storing, H. J. (Ed.). (1962). *Essays on the scientific study of politics*. New York: Holt, Rinehart & Winston.
- Storing, H. J. (Ed.). (1981). *The complete anti-Federalist* (Vols. 1–7). Chicago: University of Chicago Press.
- Storing, H. J. (1995). *Toward a more perfect union*. Washington, DC: AEI Press.
- Strauss, L. (1952). *Persecution and the art of writing*. Glencoe, IL: Free Press.
- Strauss, L. (1953). *Natural right and history*. Chicago: University of Chicago Press.
- Strauss, L. (1959). *What is political philosophy?* Chicago: University of Chicago Press.
- Strauss, L. (1962). An epilogue. In H. J. Storing (Ed.), *Essays on the scientific study of politics* (pp. 307–327). New York: Holt, Rinehart & Winston.
- Strauss, L. (1964). *The city and man*. Chicago: University of Chicago Press.
- Strauss, L. (1991). *On tyranny* (Rev. & expanded ed.). New York: Free Press.
- Strauss, L. (1997). *Jewish philosophy and the crisis of modernity*. Albany: SUNY Press.
- Tulis, J. (1987). *The rhetorical presidency*. Princeton, NJ: Princeton University Press.
- Zentner, S. (2003). The philosopher and the city: Harry Jaffa and the Straussians. *Interpretation: A Journal of Political Philosophy*, 30, 273–307.
- Zuckert, C., & Zuckert, M. (2006). *The truth about Leo Strauss*. Chicago: University of Chicago Press.

PART II

COMPARATIVE POLITICS

9

SYSTEMS THEORY AND STRUCTURAL FUNCTIONALISM

JOHN R. FISHER

Northwest Missouri State University

Although structural functionalism finds its roots much earlier than systems does theory, as researchers use it today, it is based on systems theory. Structural functionalism traces its beginnings back to the ancient Greeks and the writings of Aristotle (Susser, 1992). Systems theory emerged much later. Although the discussion of systems began with biologists in the 19th century, systems theory was not fully articulated until the 1920s. Ludwig von Bertalanffy (1956, 1962), who developed *general systems theory*, was a principal in establishing it as a field of study. Although systems theory originated later than functionalism, when researchers study functions within their structures, they do it within the scope of systems. The study of political systems came into its own with the adoption of a structural-functional approach.

The systems approach of David Easton (1965a, 1965b) and Karl W. Deutsch (1963) grew out of sociological and communication theory and a “move toward the theory and data of politics” (Almond & Powell, 1966, p. 12). Easton and Deutsch followed a communication, or cybernetic, model to study politics. Gabriel A. Almond’s study of political systems grew out of a tradition of political theory and draws from sociological and communications theories. While Easton and Deutsch adopted a purely systems approach, Almond applied structural functionalism to

systems theory. Both have value in the study of political systems.

Systems Theory

A system, according to Anatol Rapoport (1966, 1968), is a set of interrelated entities connected by behavior and history. Specifically, he stated that a system must satisfy the following criteria:

1. One can specify a set of identifiable elements.
2. Among at least some of the elements, one can specify identifiable relations.
3. Certain relations imply others.
4. A certain complex of relations at a given time implies a certain complex (or one of several possible complexes) at a later time. (Rapoport, 1966, pp. 129–130)

This definition is broad enough to include systems as different as the solar system and language. Social systems, including economics and politics, fit within the definition. Social systems might be described as a class of entities (individuals, families, institutions) with relations among them (communication channels, influence, obligations). Systems are classified by the “nature of their relation to their environments” and the “search for laws governing the

behavior of each class" (Rapoport, 1968, p. 453). Systems appear to have "a will" of their own and a "purpose" to maintain a steady state. Living systems do this through homeostasis mechanisms that restore equilibrium. Social systems have similar mechanisms (Rapoport, 1968).

While systems in the physical sciences (like the solar system, chemical reactions, and ecological systems) are extremely rigorous, social systems are less precise. In social systems, the elements and relations are vague and hard to define. As the basic unit of social systems, roles are commonly difficult to identify and classify. For the "hard" sciences, this ambiguity would be regarded as problematic, but with the social sciences, it would be commonplace (Rapoport, 1966).

The Political System

A long-standing problem of political science has been to describe and account for the internal structure of the political system. According to William Mitchell (1968), *structure* is generally applied to patterns of power and authority that characterize the relationships between the rulers and the ruled. These relationships are enduring and thus predictable.

In systems theory, the unit of analysis for these power relations is *role*, a concept developed in social psychology and applied to sociology. Political roles deal with decision making on behalf of society and with performing actions that implement the decisions and allocate scarce resources. In analyzing the political system, the researcher typically describes these roles and the people performing them. Traditionally, the main approach to classification has been "the distribution of power" (Mitchell, 1968, p. 474) among the members of the system. Because the one dimension of roles has inadequately described political systems, systems analysts have developed more inclusive variables that lend themselves better to measurement (Mitchell, 1968). Talcott Parsons (1951) put forth a set of variables that he called *pattern variables*. Gabriel Almond (1956; Almond & Coleman, 1960) suggested classifying structures based on (a) the degree of differentiation between structures, (b) the extent to which the system is "manifest" or "visible," (c) the stability of the functions of the various roles, and (d) the distribution of power. Mitchell (1968) added a fifth dimension, concerning the "sustainability of roles."

A system is generally thought of as being self-contained and distinct from its environment, with observable boundaries. In the process of determining formal members (or citizens) and their actions, boundaries are arbitrarily assigned to the political system. However, most systems are subject to external influences. Thus, analysis must also be concerned with "detecting relationships across boundaries" as inputs and outputs (Mitchell, 1968, p. 475). Yet no common language exists to describe these boundary exchanges of inputs and outputs. Easton (1957, 1965a) saw inputs as consisting of *demands* and *support* while Almond and James Coleman (1960) used the terms *political socialization*, *recruitment*,

interest articulation, *interest aggregation*, and *political communication*. Easton called the outputs *decisions*, and Almond and Coleman describe outputs as *rule making*, *rule application*, and *rule adjudication*. Mitchell (1962) used the terms *expectations and demands*, *resources*, and *support* for inputs and *social goals*, *values and costs*, and *controls* to express political outputs.

While boundary exchanges play an important part in the analysis of political systems, the main concern is with the internal processes of a system. An early area of inquiry dealt with the question of how politics would allocate scarce resources (Easton, 1953; Mitchell, 1968). Other areas of process investigation concerned the stability of systems, political socialization, and other support inputs. A third area of examination surrounded the means of ensuring loyalty and stimulating public participation. A fourth area looked at the means of achieving collective goals "from diverse individual demands" (Mitchell, 1968, p. 475). Finally, the process of dealing with problems within the political system became a matter of inspection. Mitchell (1962) viewed the internal processes of the polity as parallel to those of the larger social system. He suggested focusing on goal attainment, adaptation, system maintenance and tension management, and integration.

Applying Systems Analysis

Easton (1966) proposed to define political systems more broadly than did Rapoport. Easton defined a system as "any set of variables regardless of the degree, of inter-relationship among them" (p. 147). He preferred this definition because it freed the researcher from the need to prove that a political system is really a system. The only question of importance became whether the system was interesting and thus worth studying. The analysis need only provide understanding and an explanation of the human behavior that was of concern to the researcher.

Easton (1953, 1966) suggested that a political system was distinct from other systems because it concerned itself with "the interactions through which values are authoritatively allocated for a society" (1966, p. 147). He divided the political environment into two parts: the *intrasocietal* and the *extrasocietal*. The first comprises those systems in the same society as the political system that are not political systems because they do not have political interactions. Intrasocietal systems form the segments of society of which the political system is a component, including the economy, culture, social structure, and personalities. These systems create and shape the conditions in which the political system operates. A changing economy, culture, or social structure all have impact on political life.

The extrasocietal environment includes all the systems that are outside the given society. They may form a *suprasystem* of which the political system may be a part. An example of an extrasocietal system is the international cultural system.

From the intra- and extrasocietal systems come influences that may cause possible stress on the political system. Internal or external disturbances to the intra- and extrasocietal systems may cause stress on the political system and thus change it. However, it is also possible that some disturbances may aid in the persistence of the system while others may be neutral with regard to stress. If political systems are to continue, they must fulfill two functions. They must be able to allocate values to society and get most members of society to accept the values. The allocation of values for a society and compliance with them are essential variables of political life and distinguish political systems from other systems. By identifying these essential variables, researchers can determine when and how disturbances can cause stress to the system.

Easton (1966) provides examples of defeat at the hands of an enemy or of a severe economic crisis causing widespread disorganization and disaffection. When authorities are unable to make decisions or decisions are no longer accepted by societal members, system allocations of values are no longer possible, and the society collapses. More likely, the disruption of a political system is not that complete, and the system continues in some form. As long as the system can keep these essential variables operating, the system will persist. The capacity to counter stress is crucial to the survival of the system. The system's history of response to stress allows analysts to determine whether it is able to survive disturbances. Easton (1966) claimed that systems analysis is especially suited "for interpreting the behavior of the members in a system in the light of the consequences this behavior has for alleviating or aggravating stress upon the essential variables" (p. 149).

According to Easton (1966), systems analysis provides a way of determining the impact of the many diverse environmental influences on a system. In this way, it is possible to reduce the blow of stresses on the system and recommend appropriate action. Through the use of the concepts of inputs and outputs, the enormous variety of influences can be reduced into a manageable number of indicators. The distinction between a political system and other systems allows for interpretation of behaviors in the environment as exchanges or transactions that cross the boundaries of the political system. Easton used the term *exchanges* to refer to "the mutuality of the relationships between the political system and the other systems in the environment" (p. 150). The term *transactions* was used "to emphasize the movement of an effect in one direction, from an environmental system to the political system, or the reverse, without being concerned at the time about the reactive behavior of the other system" (p. 150).

Inputs and Outputs

Because systems are coupled together, all behavior in society is interdependent. To trace the complex exchanges and reduce them to manageable proportions, Easton condensed the main environmental influences into a few

indicators. He designated the effects that are transmitted across the boundary of a system toward some other system as the *outputs* of the first system and the *inputs* of the second system. A transaction or an exchange between systems can be viewed as a linkage between them in the form of an input-output relationship.

Inputs serve as a powerful analytic tool because they summarize variables that "concentrate and mirror everything in the environment that is relevant to political stress" (Easton, 1966, p. 150). The extent to which inputs can be used as summary variables depends on how they are defined. In their broadest sense, they include "any event external to the system that alters, modifies, or affects the system in any way" (p. 150). However, by focusing on boundary-crossing inputs dealing with the most important effects contributing to stress, one can simplify the task of analyzing the impact of the environment. Analysts no longer need "to deal with and trace out separately the consequences of each type of environmental event" (p. 150). For this purpose, Easton (1966) recommends focusing on two major inputs: demands and support. "Through them, a wide range of activities in the environment can be channeled, mirrored, summarized, and brought to bear upon political life," he wrote, and "Hence, they are key indicators of the way in which environmental influences and conditions modify and shape the operations of the political system" (p. 151). As inputs to a system, demands and supports can be of different types: material and political demands, as well as material and political supports. Easton (1965b) cites expressions of opinion and calls for a decision as examples of demands. A flood may create grievances that lead to demands for building a dam. The conventional way of making demands is to make individual requests, write letters, and carry out other forms of lobbying. More unconventional approaches to making political demands would be to demonstrate or picket. As citizens, through letters, polls, or voting, voice agreement with a decision to build the dam, they provide political support. The willingness to pay taxes to build the dam is also a form of support. Demands and supports are closely interrelated. Easton states that "by the very act of voicing a demand or proposing it for serious discussion, a member will imply that he supports it in some measure" (p. 51). By examining the changes in the inputs of demands and support, analysts can determine the effects of the environmental systems transmitted to the political system.

Similarly, outputs help interpret "the consequences flowing from the behavior of the members of the system rather than from actions in the environment" (Easton, 1966, p. 151). Since the activities of members of the system have an impact on their own subsequent actions or conditions, those actions that flow out of a system into its environment cannot be ignored. Because a great amount of activity takes place within a political system, it is useful to isolate those elements that are important in understanding the system. One way of doing this is to examine the impact of inputs (reflected as demands and support) on political

outputs. Easton defines political outputs as the decisions and actions of the authorities. A government's decision to build a dam would be a political output; the actual building of the dam would be a material output.

This approach was a departure from previous research that examined the complex political processes internal to a system in terms of who controls whom in the various decision-making processes. While the pattern of power relationships helps to determine the nature of the outputs, the outcomes of internal political processes are most useful in tracing the consequences of behavior within a political system for its environment.

Easton (1966) claimed that "outputs not only help to influence events in the broader society of which the system is a part, but also, in doing so, they help to determine each succeeding round of inputs that finds its way into the political system" (p. 152). By identifying this "feedback loop," analysts can explain the processes the system can use to cope with stress and make recommendations that alter the system's future behavior. Easton describes the feedback loop as consisting of "the production of outputs by the authorities, a response by the members of the society to these outputs, the communication of information about this response to the authorities, and finally, possible succeeding actions by the authorities" (p. 152). For actions to be taken to satisfy demands or create conditions that will do so, information must be provided to authorities (those people who speak on behalf of the system) about the effects of each round of outputs. Since a drop in support is an important source of stress, information feedback to these authorities is crucial so that they can "bolster the input of support for themselves or for the system as a whole" (p. 152). Information about the consequences of each round of outputs and about the changing conditions that impact members is essential because it enables authorities to take action to keep support at a minimal level. Appropriate response to the feedback process can have "a profound influence on the capacity of a system to cope with stress and persist" (p. 152).

Criticisms of Systems Analysis

Criticisms of systems analysis have focused mainly on three areas: methodological weaknesses of the approach, the lack of suitability for empirical research, and strong political bias (Mitchell, 1968; Susser, 1992). Some critics claim systems analysis is misleading because it assumes that "reality 'really' consists of systems." This view suggests that "societies consist of far more individual and isolated events than systems [analysis] is capable of handling" (Mitchell, 1968, p. 477). Another aspect of the criticism is that identifying boundaries and variables in the system is difficult, thus making it hard to formulate operational definitions and perform empirical research. Furthermore, critics claim that the concept of equilibrium cannot be operationally defined except perhaps in terms of economic behavior. Finally, although the inputs and

outputs can be readily identified, they may not have been adequately studied.

Bernard Susser (1992) indicated that Easton's brand of "input-output" analysis is used very little in actual research, and when it is used, "its contribution turns out to be more terminological than real" (p. 185). The problem is that it is practically impossible to study a system without looking at the past. Without understanding the system's development and its historical strengths and weaknesses, it would be difficult to tell whether an event is a crisis or a normal situation.

While systems theory generally is regarded as being supportive of the status quo and thus conservative in its nature, it is interesting to note that at the time Easton proposed systems analysis for politics, many people considered it as having a liberal bent. The 1960s was a time when behavioralists made great contributions to research in many fields. Conservatives looked at systems analysis as value-laden based on strong conceptualizations as opposed to neutral impassionate science. In addition, looking at political systems as equilibrium seeking, self-balancing entities also suggested clear ideological biases. However, systems analysis had none of the "stress, contradiction, conflict, and imbalance [that] characterize the 'normal' condition of the modern state" (Susser, 1992, p. 186) proposed by Marxists. Easton's system's "normal" state was one of "adaptive dynamic stability" (Susser, 1992, p. 186).

Structural Functionalism

The terms *functional analysis* and *structural analysis* have been applied to a great variety of approaches (Cancian, 1968; Merton, 1968). With their broad use in the social sciences has come discussion of the appropriateness of the use of structure and function and the type of analysis associated with the concepts (Levy, 1968). The functional approach is used more often than any other method in the study of Western political science (Susser, 1992). The professional literature is full of references to the "functions" of political systems and to the relation between structure and function. Sometimes the terms are used without a clear understanding of the meaning of the functionalist position, more as linguistic fashion. This section deals with the theoretical implications of structural functionalism and its relationship to political science.

Although structural functionalism predated systems theory, it still presupposes a "systems" view of the political world. Similarities link functionalism to systems analysis. Susser (1992) writes that both focus on input-output analysis, both see political systems as striving for homeostasis or equilibrium, and both consider feedback in their analysis. Yet functionalism is significantly different.

History of Structural Functionalism

Structural functionalism has a lengthy history in both the social sciences (Merton, 1968) and the biological

sciences (Woodger, 1948). Functionalism's history goes back to Aristotle's study of ultimate causes in nature or of actions in relation to their ends, or utility. Developed in 17th-century France, Montesquieu's doctrine of separation of powers is based on the notion of functions that are best undertaken separate from each other as a means of ensuring stability and security.

Functionalism became important when Darwin's evolutionary theories began to influence thinking about human behavior. Darwin conceived of the idea of survival in functional terms. Each function was important to the survival of the whole system. Systems that could not adapt their functions ceased to exist. Other students of human behavior borrowed these ideas, applying them to social affairs. Thus, social Darwinism imported these same functionalist categories into social analysis. Social Darwinists claimed that society benefited from unrestrained competition between units, that functional adaptability was required for survival, and that attempts to protect the weak hampered the functioning of society as a whole. These ideas first influenced anthropology and then sociology. Implicitly through the works of Émile Durkheim and explicitly through Parsons (1951) and Robert Merton (1968), these ideas became central to the social sciences. Almond's "Introduction" to *The Politics of Developing Areas* (Almond & Coleman, 1960), applied functionalist ideas to political life.

Susser (1992) indicates that the analogy of human social life is organic, not mechanical. Mechanical analogies imply a certain "looseness of association" (p. 203) between the parts. While the parts of a motor function as a unit, parts can be easily removed and replaced, making their union less essential and the ability to exist autonomously less likely. In the organic analogy, "Individual elements depend on the whole for their maintenance" (p. 204). Functionalists tend to view social and political units in more holistic, organic terms. "Social practices are said to have a functional role in sustaining the system as a whole" (p. 204). Functionalists equate structure to anatomy and functions to the physiology of organisms.

When only structural categories are used to make political comparisons, "The comparative analysis of political systems breaks down as the difference between compared structures increases" (Susser, 1992, p. 205). For example, the structures between a Western democracy and an African tribe are so very different as to make comparison difficult. However, functions are much more comparable. Although a prime minister and tribal chief are difficult to compare institutionally, they nevertheless serve many similar functions. Although the structures of political rule may be very dissimilar, the functions that political systems perform are universal. Although undeveloped political systems assign numerous functions to a single person or institution, in more developed political systems, the same functions may be performed by many individuals or institutions. One of the primary areas of study in functionalism is the "interplay" between the dynamic functions of a system and the more static structures it designs for itself.

Varieties of Functional Analysis

Most functional approaches share one common element: "an interest in relating one part of a society or social system to another part or to some aspect of the whole" (Cancian, 1968, p. 29). Three types of functionalism exist within this approach, and most functional analysis contains all three. The first is based on the concepts and assumptions of sociology; the second, on the supposition that social patterns maintain the larger social system; and the third, on "a model of self-regulating and equilibrating systems" (p. 29).

Francesca M. Cancian (1968) describes two distinctive types of functional analysis: traditional and formal. *Traditional functional analysis* is the most commonly used. It is based on the premise that all social patterns work to maintain the integration and adaptation of the larger system. Two attributes further distinguish traditional functional analysis from other forms of analysis. First, a social pattern is explained by the effects or consequences of that pattern, and, second, these results must be beneficial and necessary to the proper functioning of society. Researchers take one of two tacks when using traditional functional analysis. They may examine only a few aspects of society at a time and attempt to link one social pattern with one need and thus explain the pattern. Alternatively, they may deal with more complex systems, trying to show how these elements are interrelated so as to form an adaptive and consistent system.

Formal functional analysis is called *formal* because it does not include a theoretical orientation or a substantive hypothesis about events. Rather it examines the relationships between elements. It contrasts with the traditional type of analysis in that its proponents reject the attributes of "integration" and "adaptation" in favor of an examination of the equilibrating or feedback functions in systems. The effects of a trait are used to explain the system rather than the trait. No restrictions exist on the kinds of consequences that are considered. Consequences may or may not be beneficial or necessary for society.

Cancian (1968) provides an example to contrast the two types of analysis with the nonfunctionalist approach. A nonfunctionalist would explain adolescent rebellion by examining the causes of the rebellion. A traditional functionalist would explain the effects or functions of the rebellion. A formal functionalist would focus on the equilibrating or feedback systems and not on the relationships of one-way effect or cause. In practice, Cancian noted, these approaches are usually combined. Almond and Coleman (1960) rejected traditional analysis, adopting a more formal approach.

Applying Functional Analysis to the Study of Politics

According to Michael G. Smith (1966), four approaches are useful in the comparative study of political systems: process, content, function, and form. Studies based on

process and content face huge obstacles. In developed countries, the processes of government are “elaborately differentiated, discrete and easy to identify,” but in simpler societies, the same processes are “rarely differentiated and discrete” (p. 114). They occur within the context of institutional activities that are difficult to analyze for political processes. The more “differentiated and complex” the government processes, the “greater the range and complexity” (p. 114) of content. Since content and process are “interdependent and derivative,” they require independent criteria for studying government (p. 114).

The functional approach does not have the same limitations as process and content. It defines government as all those activities that influence “the way in which authoritative decisions are formulated and executed for a society” (Easton, 1957, p. 384). From this definition, various schemata were developed to study the functions of government. Easton listed five modes of action as elements of all political systems: legislation, administration, adjudication, the development of demands, and the development of support and solidarity. These were grouped as input and output requirements of political systems. According to Almond and Coleman (1960), the required inputs are political socialization and recruitment, interest articulation, interest aggregation, and political communication. As outputs, he identified rule making, rule application, and rule adjudication.

In 1960, Almond and Coleman were the first to compare the political systems of “developing” areas systematically according to a common set of categories. To do this, they felt, they could no longer rely on the comparative approaches used to study governments in Western Europe. To find concepts and categories appropriate for use in comparing developing countries, they turned to sociological and anthropological theory (Almond & Coleman, 1960). Rather than adding new terms, they adopted and adapted an old vocabulary to a new situation. Instead of the concept of *state*, which would be limited by legal and institutional meanings, they used *political system*; instead of *powers*, with its legal connotations, they preferred *functions*; instead of *offices*, they used *roles*; instead of *institutions*, which directs thinking toward formal norms, they used *structures*; and instead of *public opinion* and *citizenship training*, they preferred *political culture* and *political socialization*.

In order to develop a system of categorization for all societies, regardless of size and culture, Almond and Coleman (1960) had to modify their definitions of politics and political systems. They felt the definitions of politics that identified societal functions as integration and adaptation were inadequate in describing their concept of political systems. Instead, they borrowed from Max Weber’s concept of state and Easton’s view of power. Easton (1953) offered a definition with three components: “The political system allocates values by means of policies; the allocations are authoritative; and its authoritative

allocations are binding on society as a whole” (p. 130). Almond and Coleman (1960) sharpened Easton’s definition of authority by building in Weber’s notion of *legitimate physical compulsion*. They viewed the political system as “the legitimate, order-maintaining or transforming system in society” (p. 7).

With the concepts of input and output, Almond and Coleman (1960) moved from a definition of *political* to that of *system*. They saw in the notion of *system* properties that interpret interactions of society, whereas *political* separated out the interactions in order to relate them to other concepts. Among the properties were comprehensiveness, interdependence, and the existence of boundaries. Systems analysis was comprehensive because it included all interactions, both inputs and outputs. It was interdependent because change in one subset of interactions would change others. The political system has boundaries in that there are points where it begins and points where it ends and other systems take over.

Political systems have common properties, according to Almond and Coleman (1960). First, all political systems, even the simplest, have political structure. Second, the same functions are performed in all political systems. Third, all political structure is multifunctional, whether in primitive or in modern societies. Finally, all political systems are “mixed” systems in the cultural sense. No society is strictly modern or only primitive.

As stated previously, Almond and Coleman (1960) listed seven functions of all political systems: political socialization, interest articulation, interest aggregation, political communication, rule making, rule application, and rule adjudication. The first four belong to the input side of a system’s functioning, and the last three to its policy outputs. Political communication links inputs to outputs in a way that provides the function of a feedback loop. Whereas Easton’s systems analysis deals primarily with “demands and supports,” Almond and Coleman’s categorization of inputs and outputs in the political system is much more extensive and in fact has led to a multifaceted approach to the study of politics.

In their study of political systems, Almond and Powell (1966) considered the activities or functions from three points of view: the conversion functions of interest articulation, interest aggregation, political communication, rule making, rule application, and rule adjudication; the operation and capabilities of the political system in its environments; and the way in which political systems maintain or adapt themselves to pressures for change over the long term. These latter functions referred to the maintenance and adaptation functions of political recruitment and political socialization.

An Example of the Functional Approach

Many of Almond and Coleman’s (1960) categories have become unique fields of study. For example, Fisher’s

research on mass media's effect on political decision making drew on Almond and Coleman's categories and mass media functions to develop a taxonomy of media functions in policy making (Fisher, 1991; Fisher & Soemarsono, 2008). Whereas the systems view often refers to the "non-descript conversion process" (Susser, 1992, p. 206), the functionalist approach deals explicitly with the steps involved from articulating requirements to fulfilling political outputs.

To show how structural functionalism fits within systems theory, Fisher's studies of mass media functions in policy making are examined (Fisher, 1991; Fisher & Soemarsono, 2008). Those studies found 14 media functions within six policy stages (Almond & Powell, 1966; Dunn, 1981; Jones, 1977; Wirt & Mitchell, 1982). To arrive at the 14 media functions in the policy process, Fisher adapted Lambeth's (1978; see also Fico, 1984) 10 media functions. Within Stage 1, problem identification and articulation, were found two media functions: (1) identification of problems by media and (2) relaying of problems to the public. Within Stage 2, policy recommendation and aggregation, the media were found to function in three ways: (3) identification of groups and proposals, (4) identification of policymaker proposals, and (5) media suggestions of content. In Stage 3, policy decision and adoption, the media functioned by (6) setting the tempo of decision making, (7) recommending how to vote, and (8) informing the public of content. Within Stage 4, policy implementation, the media functioned by (9) describing administration and (10) alerting the public to problems. Within Stage 5, policy evaluation, were found the media functions of (11) evaluating effectiveness and (12) reacting to policy. Finally, within Stage 6, policy resolution or change, were found the media functions of (13) stimulating review and (14) proposing change or termination.

In his study of lawmakers' use of reporters, Lambeth found that reporters were more influential in the five functions involving their potential impact in transmitting information to the public than in the functions involving personal or professional influence in the legislative setting. Fisher (1991; Fisher & Soemarsono, 2008) used content analysis in his study of mass media functions to determine the role of the media in informing or persuading the public and policymakers. Fisher confirmed Lambeth's finding that reporters are more influential in functions involving transmittal of information to the public and less important in functions involving personal and professional influence in the legislative setting. In addition, the study seemed to bear out Lambeth's conclusions that the impact of the press on elected officials is low to moderate.

Fisher (1991) provides an example of the relationship among systems, structures, and functions. While the policy stages are functions in the political system, they also provide structure for the media functions. The first two provide input functions in the political system. The next is a process function. The last three serve as output functions.

Terminology Used in Structural-Functional Analysis

Structural-functional analysis is made more difficult because of the confusion of terms. The difficulty in speaking about structural functionalism comes from five sources, according to Levy (1968). First, the feeling exists that structural-functional analysis is something new, when in fact it is as old as the scientific method. Second, definitions are messy because terms are unclear and refer to more than one thing. Third, many researchers make the mistake of believing that final causes can be found from their work. They assume that it is possible to find the purpose and design of the phenomena they study. This is a fallacy called *teleology*. Fourth, researchers assume that the methodology is tried and proven, when in fact models of analysis are often misunderstood and misconstrued. Finally, researchers have allowed bias to seep into their work. Unintentionally they have written evaluative approaches into their analysis, thus raising questions about objectivity.

Structural functionalism is a synonym for *scientific analysis* in general and as such has existed long before the adoption of the name *structural functionalism* in the social sciences. In the biological sciences, for example, the study of structure and function has a long history. Structural functionalism analysis consists of nothing more than stating empirical questions in one of the following forms or some combination of them: (a) What observable uniformities (or patterns) exist in the phenomenon under study? (b) What conditions result because of the phenomenon? (c) What processes occur as a result of the conditions? The first question asks: What structures are involved? The second: What functions have resulted because of the structures? Asked in the opposite direction, different results could occur: What functions exist? What structures result from the functions?

Function and Structure

Another problem, according to Levy (1968), is that the general concept of structure has many different referents, in both the biological and the social sciences. Joseph Woodger (1948) in biology and Merton (1968) in the social sciences have pointed to the abundance of referents given to the term *function*. This has led to a lot of confusion. Much of the literature is preoccupied with function, whereas structure has been discussed less. *Function* may be defined as any condition or state of affairs resulting from an operation of a unit of the type under consideration in terms of structure. In the biological sense, the unit is an organism or subsystem of an organism. In the social sciences, the unit is usually a set of one or more persons (actors). *Structure* may be defined as pattern or observable uniformity in terms of the action or operation taking place. In the social sciences, the focus of analysis has been on the

structure of societies and other social systems or the structures (patterns) of actions in general.

Classification of functions or structures depends partly on point of view. What is function from one point of view may be structure from another. Levy (1968) gave examples of this confusion. The manufacture of automobiles is production from the point of view of the automobile user but consumption from the point of view of the steelmaker. Functions in this sense are patterns or structures or have important structured (patterned) aspects, and all structures are the results of operations in terms of other structures, so they are in fact functions. The politeness of children may be considered a structure of their behavior or a function in terms of the structures (patterns) of parenting.

Requisites and Prerequisites

Functional and structural requisites are useful in the analysis of any unit. A *functional requisite* may be defined as “a generalized condition necessary for the maintenance of the type of unit under consideration” (Levy, 1968, p. 23). Functional requisites respond to the question: What must be done to maintain the system at the level under consideration? A functional requisite exists if its removal (or absence) results in the dissolution of the unit or the change of one of its structural elements.

A *structural requisite* may be defined as a pattern of action (or operation) necessary for the continued existence of the unit (Levy, 1968). To discover structural requisites, ask: What structures must be present so that operations will result in the functional requisites for the unit? Functional requisites answer the question: What must be done? Structural requisites are answers to the question: How must what must be done be done?

According to Levy (1968), *structural functional requisite analysis* includes the following steps: (a) Define the unit of phenomena to be studied, (b) discover the setting, (c) discover the general conditions (or functional requisites) that must be met if the unit is to persist in its setting with change or alteration of structures, and (d) discover what structures must be present to maintain the system.

Functional and structural prerequisites must preexist if a unit is to come into existence. Sometimes the requisites and prerequisites may be similar or identical. On the other hand, the requisites and prerequisites may not coincide. For example, the structures that must be maintained in order for the United States to continue as a highly modernized society are not the same as those that have to pre-exist for Nigeria to become highly modernized. However, the structures may be similar if one looks at the United States at the beginning of the 19th century (Levy, 1968).

Concrete and Analytic Structures

Failure to distinguish between concrete and analytic structures may result in the *fallacy of reification* (or

misplaced concreteness). For example, the terms *economy* and *polity* cannot occupy the same position in system analysis as the term *family*. Family is an example of a concrete structure, as are business firms, governments, and societies. In *concrete structures*, the units are capable of physical separation from other units of the same sort, and membership is easily defined. In *analytic structures*, no concrete separation of units is possible. For example, no social system is without economic and political structures (Levy, 1968).

Institutions, Traditional Structures, and Utopian Structures

Although these terms are sometimes used interchangeably, they refer to different types of structures. Institutions are structures with normative patterns with which conformity is expected, and failure to conform is sanctioned or met with indignation. The structure becomes a requisite of the system. The structure does not change without destroying the structural requisite. For example, age and role are tied together in all societies. If the requisite age changes for certain roles or functions, the structure would also change.

Traditions are *institutionalized* as the structure is perpetuated to the extent that changes in functions do not have an effect on the structure. Tradition is a *double institution*, according to Levy (1968): “The structure concerned is an institution and the perpetuation of the structure is also an institution” (p. 27). Important traditions may vary in conformity and sanctions. The tradition of driving on the right-hand side of the road would not have the same level of sanctions as the tradition against incest.

Utopian structures, although they may not be institutionalized, still require adherence as institutional ideals (Levy, 1968). The principle “Love thy neighbor as thyself” is an ideal that is institutionalized in some social contexts. Its perpetuation is also institutionalized. Utopian structures allow the teaching of societal norms and the perpetuation of structures.

Ideal and Actual Structures

Members of a society establish ideal structures to determine how they should behave, whereas actual structures are patterns of how they do behave. Although sometimes the ideal and the actual coincide, more often they do not fit perfectly. This difference in fit causes stress in the social system. Only with perfect knowledge and perfect motivation would there be a perfect fit between the ideal and the actual structures.

Criticisms of Structural Functionalism

Critics of structural functionalism view it as “a translation of Anglo-American political norms in methodological

terminology” (Susser, 1992, p. 207). Structural functionalism may be in decline as a methodological approach for the study of politics; however, it leaves a set of terms that are still used in political jargon. Some of those in the functionalist camp (Merton among them) rejected the notion of this decline. “Much of what was best in the political research of an entire generation was couched in its terms” (Susser, 1992, p. 207).

One of the main criticisms of structural functionalism is that its categories were “too undifferentiated to be of real help in actual research” (Susser, 1992, p. 206). Although Almond’s functional taxonomy has greater specificity and serviceability than the systems approach, it is seen as not much more than a translation of familiar and known phenomena into blandly broad categories. As such it promotes “a terminological rather than an essential transformation in the discipline” (Susser, 1992, p. 206).

Another criticism is related to the methodological approach used in functionalism. A list of functions is created deductively and then appropriate structures are identified. In some cases, this approach leads to “empirical contortions” to satisfy the framework. This criticism applies to much academic research, leaving the researcher, rather than the approach, responsible for assuring research validity.

A final criticism, according to Susser (1992), is that functionalism “harbors an ideological slant” (p. 207) that sustains existing structures. It describes what exists rather than what ought to be, thus maintaining the status quo.

As if anticipating this criticism, Almond and Powell (1966) responded to the criticism that functional-systems theories imply “an equilibrium or harmony of parts” and “that they have a static or conservative bias” (p. 12). Political systems are not necessarily harmonious or stable, they wrote, but interdependent. The task of political science research is “to ascertain how change in any one of the parts of a political system affects other parts and the whole” (p. 13). They built political development into their approach to the study of systems. They look at political systems “as whole entities shaping and being shaped by their environments” (p. 14). To understand the processes of political development, they examine the interaction of the political system with its domestic and international environments.

Conclusion

The study of structural functionalism and systems theory had its heyday with the works of Easton (who examined political systems), Merton (noted for his study of social structure), and Almond and Coleman (who developed a taxonomy of political functions within political systems). A majority of political studies from that period used systems theory and structural functionalism as their framework (Susser, 1992). While few researchers today claim a framework based on

these theories, the approach is still alive and well (Charnock, 2009; Fisher & Soemarsono, 2008; Fisk & Malamud, 2009; Mohamed, 2007; Scheuerell, 2008). Understanding politics requires political syntax, much of which continues to be based on structural functionalism and systems theory.

References and Further Readings

- Almond, G. A. (1956). Comparative political systems. *Journal of Politics*, 18, 391–409.
- Almond, G. A., & Coleman, J. S. (1960). *The politics of the developing areas*. Princeton, NJ: Princeton University Press.
- Almond, G. A., & Powell, G. B., Jr. (1966). *Comparative politics: A developmental approach*. Boston: Little, Brown.
- Bertalanffy, L. von. (1956). General system theory. *General Systems*, 1, 1–10.
- Bertalanffy, L. von. (1962). General system theory: A critical review. *General Systems*, 7, 1–20.
- Cancian, F. M. (1968). Varieties of functional analysis. In D. L. Sills (Ed.), *International encyclopedia of social sciences* (Vol. 5, pp. 29–41). New York: Macmillan.
- Charnock, G. (2009, Summer). Why do institutions matter? Global competitiveness and the politics of policies in Latin America. *Capital & Class*, 33(98), 67. Retrieved August 5, 2009, from <http://find.galegroup.com/gtx/start.do?prodId=AONE>
- Coser, L. A. (1975). *The idea of social structure: Papers in honor of Robert K. Merton*. New York: Harcourt Brace Jovanovich.
- Deutsch, K. W. (1963). *The nerves of government*. New York: Free Press.
- Dunn, W. N. (1981). *Public policy analysis: An introduction*. Englewood Cliffs, NJ: Prentice Hall.
- Easton, D. (1953). *The political system: An inquiry into the state of political science*. Chicago: University of Chicago Press.
- Easton, D. (1957). An approach to the analysis of political systems. *World Politics*, 9, 383–400.
- Easton, D. (1965a). *A framework for political analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Easton, D. (1965b). *A systems analysis of political life*. New York: Wiley.
- Easton, D. (1966). Categories for the systems analysis of politics. In D. Easton (Ed.), *Varieties of political theory* (pp. 143–154). Englewood Cliffs, NJ: Prentice Hall.
- Fico, F. (1984). How lawmakers use reporters: Differences in specialization and goals. *Journalism Quarterly*, 61(4), 793–800, 821.
- Fisher, J. R. (1991). News media functions in policymaking. *Canadian Journal of Communications*, 16(1), 139–145. Retrieved April 3, 2009, from <http://www.cjc-online.ca/index.php/journal/article/view/589/495>
- Fisher, J. R., & Soemarsono, A. (2008). Mass media impact on post secondary policy making: A case study of a failed merger attempt. *Competition Forum*, 6(1), 96–102.
- Fisk, C. L., & Malamud, D. C. (2009). The NLRB in administrative law exile: Problems with its structure and function and suggestions for reform. *Duke Law Journal*, 58(8), 2013–2085. Retrieved August 5, 2009, from <http://find.galegroup.com/gtx/start.do?prodId=AONE>

80 • COMPARATIVE POLITICS

- Jones, C. O. (1977). *An introduction to the study of public policy* (2nd ed.). Boston: Duxbury.
- Lambeth, E. B. (1978). Perceived influence of the press on energy policy making. *Journalism Quarterly*, 55(1), 11-18, 72.
- Levy, M. J. (1968). Structural functional analysis. In D. L. Sills (Ed.), *International encyclopedia of the social sciences* (Vol. 5, pp. 21-29). New York: Macmillan.
- Merton, R. K. (1968). *Social theory and social structure*. New York: Free Press.
- Mitchell, W. C. (1962). *The American polity: A social and cultural interpretation*. New York: Free Press.
- Mitchell, W. C. (1968). Political systems. In D. L. Sills (Ed.), *International encyclopedia of the social sciences* (Vol. 5, pp. 473-479). New York: Macmillan.
- Mohamed, J. (2007). Kinship and contract in Somali politics. *Africa*, 77(2), 226-249.
- Parsons, T. (1951). *The social system*. Glencoe, IL: Free Press.
- Rapoport, A. (1966). Some system approaches to political theory. In D. Easton (Ed.), *Varieties of political theory* (pp. 129-142). Englewood Cliffs, NJ: Prentice Hall.
- Rapoport, A. (1968). General systems theory. In D. L. Sills (Ed.), *International encyclopedia of the social sciences* (Vol. 5, pp. 452-457). New York: Macmillan.
- Scheuerell, S. K. (2008). Using the Internet to learn about the influence of money in politics. *Social Education*, 72(3), 152-155.
- Smith, M. G. (1966). A structural approach to comparative politics. In D. Easton (Ed.), *Varieties of political theory* (pp. 113-128). Englewood Cliffs, NJ: Prentice Hall.
- Susser, B. (1992). *Approaches to the study of politics*. New York: Macmillan.
- Wirt, F. M., & Mitchell, D. E. (1982). Social science and educational reform: The political uses of social research. *Educational Administration Quarterly*, 18(4), 1-16.
- Woodger, J. H. (1948). *Biological principles: A critical study*. London: Routledge.

10

POLITICAL DEVELOPMENT AND MODERNIZATION

SARAH WILSON SOKHEY

The Ohio State University

To eat and to talk—to be free from hunger and repression: These elementary values animate a worldwide quest for political democracy and economic rationality.

Adam Przeworski (1991)

Political science has long been concerned with how to establish systems allowing us “to be free from hunger and repression.” *Political development* implies that some governments are better at accomplishing these goals than others are. Although we should be careful not to idealize democracy with all its imperfections—and indeed Samuel Huntington would remind us that political order matters more—many agree that democracy in some form is preferable to the wide array of nondemocratic systems of government. *Modernization* refers to economic development and the transformation from agricultural to industrial societies, along with corresponding social and cultural shifts (although the use of terms such as *modern* and *primitive* has been criticized as inappropriately stereotyping certain cultures from a Western perspective).

The central question of how economic conditions are linked with the emergence of democracy or dictatorship has been a topic of interest from the time of ancient scholars through contemporary political science. Aristotle noted that

democracy could not function well in a society in which a large proportion of the population lived in poverty.¹ In a study of early American democracy, Tocqueville (1835) also noted that democratic systems would suffer in societies with great economic inequality; where inequality and democracy coexisted, class cleavages would define politics and the poor would vote to redistribute wealth from the rich.² Reacting to the spread of Communism around the world beginning in the 1950s, scholars and politicians in democratic countries concerned themselves with the necessary prerequisites of a democratic society, including economic factors, in order to predict which countries were likely to become or stay democratic. On the collapse of many communist regimes in the early 1990s, our attention again turned to explaining the relationship between economics and politics as countries struggled to transition to democracy and to market-based economies simultaneously. Questions about the prerequisites for successful democracy have continued to be relevant more recently in countries such as Iraq and Afghanistan.

This chapter traces the study of political development and modernization. First, it discusses the origins and development of *modernization theory*, which encompasses a set of explanations linking economic, social, and cultural changes with shifts in political systems. Modernization theory is a starting point for understanding how contemporary political scientists approach this topic. It puts forth the notion that economic development leads to social and cultural changes that alter the political behavior of citizens and ultimately result in democratic governments. Second, the chapter turns to the empirical evidence supporting modernization theory and critiques of the theory's larger applicability around the world. Critics suggest that some types of economic development might actually prove to be destabilizing, rather than advancing the social and cultural elements that provide the foundation for democratic societies. Others have suggested that although wealth does not explain the emergence of democracy, the odds of a country's remaining democratic are higher in richer countries. Third, this chapter reviews the policy implications of what scholars have learned about the connection between economic and political changes. Finally, it looks to future directions in research on political development and modernization.

Modernization Theory

Modernization theory refers to a group of explanations linking economic development and accompanying social transformations with the type of political system that emerges. The basic story of modernization theory is as follows: As countries modernize economically, they transition from agricultural to industrial societies. Industrialization results in urbanization, which means that more of the population lives in cities in which they have greater access to information, media, and education. Increasing economic wealth, which accompanies industrialization, results in a growing middle class that begins to participate more in politics and make demands on the government. Ultimately, the resulting changes in mass political behavior make the emergence and survival of democratic governments more likely. The emphasis in this branch of research is on the stages of development from more traditional to more advanced societies. Countries are all on a similar path, but some countries are further along than others, and ultimately we should expect to see all countries develop democratic systems, albeit at different rates. For instance, in *The System of Modern Societies*, Talcott Parsons (1971) examines the development of the state in the context of Western Europe, detailing how societies evolve from traditional to modern ones.

Critically, these early sociological explanations of modernization were based almost exclusively on the Western European context. As economic development progressed according to Western trends such as industrialization, politics would change as well. Karl Deutsch

(1953, 1961) suggested that socioeconomic development alters mass behavior in politics. In doing so, Deutsch developed the concept of *social mobilization*, a critical component of the modernization process. Social mobilization denotes changes happening in a large portion of a society as it transitions from being traditional to being modern. For instance, as countries industrialize and rely less on agricultural production, urbanization provides individuals and groups with the exposure to information and the resources with which to participate in politics. Urbanization, education, and the development of new social networks and roles in society are all part of social mobilization, which, ultimately, leads individuals to make new demands on their governments. Deutsch notes, however, that the democratic participation resulting from social mobilization might manifest as communal riots and civil wars, not the peaceful types of participation, such as voting, that we typically associate with democracies today. In whatever form it takes, mass political participation should threaten the continued existence of governments that do not respond to citizens, thereby paving the way for more democratic societies.

In keeping with the important role of social mobilization in development, the effects of modernization alter how citizens view themselves in connection with the state. In *The Passing of the Traditional Society*, Daniel Lerner (1958/2000) uses a case study of a Turkish village to examine the effects of the transition from a "traditional" to a "modern" society. Lerner concludes that modernization—through exposure to mass communication, increasing literacy, and new transportation networks—profoundly alters the way that individuals see themselves in the state. Individuals go from seeing themselves as subjects of the state to citizens who have the leverage to make demands on their rulers.

Seymour Martin Lipset (1959, 1960) argues that modernization is likely to spur the establishment of democracy as well as its survival. As countries develop economically, they also become more complex, and in particular, the development of a stronger middle class will allow for the development of a more politically active civil society. Lipset notes that education is likely to encourage the development of citizens who are aware of their government and are more likely to participate by making demands. Even in developed countries, we know that more educated citizens are more likely to participate (although the effect of individual education and wealth on political participation varies by country). Through this process, governments will be forced to make concessions to the empowered and increasingly politically aware middle class, or the regimes will fall due to uprisings. These concessions will eventually produce democracies.

The middle class was thought to be a critical component of the modernization process such that, in explaining the origins of democracy, Barrington Moore (1966) concludes, "no bourgeois, no democracy" (p. 418). One example of the critical role played by the middle class can be seen in

the English state-building process. The development of wool manufacturing in England created an increasingly wealthy middle class. Sheep were easier to hide from the state than resources such as land, making the taxation of the wool industry difficult. To successfully collect revenue, the English state was forced to make concessions to the middle class, which in turn spurred the development of a different type of citizen–state relations. Thus, the development of an empowered middle class was the beginning of a more democratic basis of government in which citizens had more influence.

Dietrich Rueschemeyer, Evelyn Stephens, and John Stephens (1992) offer a twist on class-based explanations, arguing that it is not an expansion of the middle class, but rather a shift in the balance of power among the classes, that explains democratization. When economic changes weaken the powerful upper class in relation to subordinate classes, democratization will occur. In a notable departure from traditional modernization theory, they conclude that the middle class often opposed further democratization after having gained some initial concessions from the state. Carles Boix (2003) also argues that it is not the growth of the middle class alone, but the balance of power among social classes, the nature of economic resources, and the distribution of wealth that influence the formation of democracy. Nonetheless, even by these different accounts, class politics and the middle class play a central, if not exclusive, role in changes in the political system.

Building on modernization theory, scholars have also suggested that democracies emerged and succeeded when certain norms and beliefs were present. The idea that certain norms are necessary for democracies to function well suggests an important qualification to modernization theory. If socioeconomic changes do not alter the beliefs and norms of citizens, then we should not expect modernization to increase the likelihood that a successful democracy will be established.

In seeking to understand the cultural basis of democracy, political scientists began to bring survey research into the discipline. *The Civic Culture*, by Gabriel Almond and Sidney Verba (1963), is a seminal study in comparative political behavior and relies on survey research. They argue that a certain set of beliefs makes it likely that a country will be able to establish a well-functioning democratic system of government. According to Almond and Verba, democracies flourish in pluralistic cultures based on communication and persuasion, consensus building, diversity, and representation in governing bodies. They identify three types of citizens—parochials, subjects, and participants. *Parochials* are those who are completely politically unaware; *subjects* are politically aware and are subservient to the state but do not participate; and *participants* are politically aware and actively engage in the political system through activities such as voting. Democracies have the largest number of participants. The least developed systems have many parochials, as in a feudal system or very poor, nondemocratic countries in

which citizens have little awareness or knowledge of the government. Other nondemocratic systems, such as communist ones, may ensure that citizens are aware of the state's presence and power (thereby fostering subjects) but do not allow for the meaningful participation that is associated with participant citizens.

Almond and Verba support their claims using survey evidence from the United States, the United Kingdom, Germany, Italy, and Mexico (ranked from the most to the least democratic country). For instance, they explain differences in the histories of British and German democracies by pointing to the presence of different types of citizens. Both countries had histories of deference to political leaders, but the United Kingdom became the epitome of democracy, whereas Germany saw the rise of a fascist political system with the Nazis. Almond and Verba's argument was that the British state did not have a history of maintaining the same exhaustive power over its people as existed in Germany; as a result, British citizens were more resistant to fascist measures than Germans were. Nearly 30 years after *The Civic Culture*, Ronald Inglehart (1990) again conducted surveys to establish the underlying cultural aspects of democracy. Using survey evidence from 25 industrial nations in the 1980s, he found that certain social characteristics, such as high levels of interpersonal trust and support for gradual change in society, are correlated with more stable democracies.

Survey research such as that done by Almond, Verba, and Inglehart provided the basis for broader studies of political culture, including the World Values Survey at the University of Michigan, which tracks opinions on social and political questions in more than 80 countries around the world. However, there is an important question of causality in applying such studies to understanding the origins of democratic societies. We do not know whether these values produce democratic societies or whether democratic societies foster citizens with these values. Consider again the example of the United Kingdom's remaining a democratic country while Hitler's fascist regime was able to come to power in Germany. Almond and Verba argue that the civic culture was stronger in the United Kingdom, thereby precluding fascist developments. However, we do not know what determined the type of citizen in the first place. Political institutions are as likely to shape citizens as citizens are to shape their institutions. If the type of citizen determines the nature of the political system and the nature of the political system shapes citizens, then this explanation is circular and unhelpful.

Despite the difficulties of assessing causality from this survey evidence, such research suggests an important qualification to modernization theory by identifying the presence of different political beliefs in different systems. This research suggests that if one wants to promote successful democratic societies, one should pay attention not only to economic growth but also to the promotion of certain norms and beliefs.

Applications, Empirical Evidence, and Critiques

Modernization theory has faced several serious criticisms. The terminology used in modernization theory has been criticized in part because it assumed that *lesser* developed countries were simply more *primitive* than countries such as the United States and those in Western Europe, implying that lesser developed countries should end up looking like Western governments. Substantive objections have been raised as well, including dramatically different explanations of the connection between economic and political developments. The more enduring criticisms, such as the destabilizing effects of modernization, emerged from a consideration of a wider range of cases. Other counterexplanations, such as *dependency theory* (discussed below), have now been rejected because they were largely based on the experiences of a few Latin American countries during a specific era.

Political Order in Changing Societies

One of the major challenges to modernization theory was posed by Huntington (1968) in *Political Order in Changing Societies*. Huntington noted that economic modernization and political development are not synonymous but rather are distinct processes. Furthermore, economic development and the rapid social changes accompanying it are as likely to result in the political decay of societies as in their development. Instability is most likely to occur in the early stages of modernization, or when there is growth followed by sudden setbacks. If socioeconomic changes lead to rapid social mobilization that outpaces the development of political institutions, then decay will certainly occur. Huntington went so far as to suggest that political order trumped other concerns, producing the controversial view that stable authoritarian systems would be preferable to unstable democratic ones.

By some accounts, Huntington “killed off modernization theory” by highlighting the destabilizing effects of rapid development.³ However, although Huntington’s work created debate about the consequences of economic development, his explanation rested on an understanding of the link between social mobilization and economic development—two factors central to modernization theory. Therefore, although Huntington created a significant challenge to the then existing understanding of modernization theory, some scholars would nonetheless classify him as a *modernizationist*, albeit one who emphasizes the importance of political institutions.⁴

Dependency Theory

In economics and the social sciences, modernization theory was originally developed through a comparison of Western European development with the less economically developed countries of Africa, Asia, and Latin America.

Economists and economic historians such as W. W. Rostow (1959) concluded that all countries would eventually reach a more advanced stage of development similar to that of the United States and countries in Western Europe. However, out of the Latin American context came one of the greatest challenges to modernization theory.

Dependency theory directly countered modernization theory, arguing that economic development might hinder social modernization and the emergence of democracy (but for reasons very different from those raised by Huntington). Dependency and modernization theory shared a similar topic of interest—why some states were developed while others lagged—but differed in their assumptions and approaches. While modernization theory focused on individual-level factors such as political behavior, dependency theory maintained that development could be explained only by considering a country’s historical role in the global political and economic system. In this sense, dependency theory represents a difference in approach (i.e., the factors it considers most relevant) rather than a distinct theory of global development.⁵

Support for dependency theory came from scholars studying Latin American countries who noted a cycle of weak countries being exploited by powerful countries in the international economy. In the 1960s and 1970s, the experiences of Latin American countries suggested that economic modernization—through integration into the world economy—might not be enough to produce stable growth or democratic systems. Scholars argued that, not only did development *not* promote democratization, but globalization put small states at an economic disadvantage in the larger international economic systems. This meant that developing countries would never experience stable economic growth because they were trapped in an inescapable dependence on wealthier nations.

In “The Development of Underdevelopment,” Andre Gunder Frank (1970) cites Chile and Brazil as examples of poor countries trapped in a cycle of economic growth that benefits richer countries but not the satellite territories that are the origins of production. This pattern of development began in the 16th century with Spanish conquest of the modern-day Chilean territory and Portuguese conquest of Brazil. While both countries went through periods of what appeared to be successful economic growth, their ultimate success hinged on how they fit into the larger international economy and whether there was foreign and domestic interest in the production of local economies. When international interest and subsequent financial investment waned, local economies foundered.

Today, dependency theory has been discredited because of a lack of evidence supporting its central arguments in other regions of the world or over long periods of time. Evidence refuting dependency theory emerged in the late 1970s, when Latin American economies in countries such as Argentina began to grow rapidly, breaking from the cycle of poverty on which dependency theory focused. Because dependency theory had predicted continued stagnation

where growth emerged, it was discounted as a credible explanation of the connection between political and economic development.⁶

Natural Resource Curse and Modernization Theory

Another shortcoming of modernization theory was its inability to explain the persistence of countries that were wealthy but not democratic. For instance, scholars refer to a *natural resource curse* in which countries with large amounts of natural resources (and which rely on those resources for a large proportion of state revenue) have difficulty in attaining stable economic growth or maintaining democratic systems of government.

In the title of her book, Terry Lynn Karl (1997) refers to the natural resource curse as “the paradox of plenty” faced by oil-rich countries. She studies several oil-rich countries—Venezuela, Iran, Nigeria, Algeria, and Indonesia—which are diverse on an array of political and social characteristics but share striking similarities resulting from their oil wealth. In these oil-rich states, politicians did not have incentives to think about the long-term economic efficiency of how they generated state revenue and distributed resources because there were strong incentives to give undue favoritism to the oil industry. The presence of large quantities of oil promoted political and economic systems that gave undue advantages to the oil industry at the cost of long-term economic efficiency (by investing more broadly in the economy). This meant that oil wealth promoted the development of centralized states, with the result that democratic processes, including widespread societal participation in politics, are unlikely to emerge. Karl refers to this problem as a kind of modern-day Midas touch, in reference to the mythical king who wished, at his own peril, that everything he touched would turn to gold.

The central risk factor for countries with the natural resource curse is that the economy relies primarily on a single commodity (or good), and this applies not just to oil. For example, Michael Shafer (1994) documents the problems posed by economies dominated by a single sector, including copper in Zambia, tea in Sri Lanka, coffee in Costa Rica, and light manufacturing in South Korea. Nor is the problem posed by natural resource wealth limited to a handful of countries. Some 75% of states in sub-Saharan Africa and more than two thirds of the countries in Latin America, the Caribbean, North Africa, and the Middle East depend on primary commodities for a minimum of half of their income from exports (Ross, 1999).

In relation to modernization theory, the existence of wealthy countries with stable dictatorships (or weak democracies) is very surprising. These oil-rich countries have proved to be stable nondemocratic systems, challenging the notion that wealth breeds a demand for rights and goods from the government that results in democracy. If modernization theory is correct, we should expect economic wealth to be associated with social and cultural

shifts that promote democratic systems of government. Thus, the existence of wealthy dictatorships—and the natural resource curse—is unexpected from the perspective of classic modernization theory. In particular, in keeping with modernization theory, we would expect the citizens of wealthy countries to become more educated, participate more in their own governance, and ultimately create the pressure for a more democratic system of governance in the world.

To explain this seemingly strange outcome, we must consider the consequences of economic growth from natural resources, in which Karl’s explanation of the paradox of plenty offers some insight. If Karl is correct that oil shapes the centralization of the state and creates incentives for the state to offer undue influence and privileges to the oil industry, this suggests that oil wealth does not promote the type of overall, long-term economic development that promotes democratization. Even though education may become more widespread and citizens may become increasingly exposed to the outside world, the state will maintain highly centralized power and will be dominated by one powerful interest group—the oil industry—such that democratic participation will be limited.

Consider the example of Iran. Iranian citizens have increasingly significant exposure to the outside world, particularly through the availability of satellite television offering antiregime perspectives and foreign news sources such as the Cable News Network (CNN) and the British Broadcasting Corporation (BBC). Dissent can also be seen in the emergence of independent newspapers and increasing Internet access in Iran. While the growth of these media outlets has prompted the state to react by expanding its own presence in satellite television, significant competing sources of information still exist.⁷ As such, modernization theory would predict that Iranian citizens should become increasingly interested in being politically active. Nonetheless, an Iranian democracy is not likely to emerge in the near future.

According to Karl’s argument, an important part of the reason Iran has not, and likely will not, become more democratic is that its reliance on oil for wealth has encouraged the development and persistence of state structures that are highly centralized and favor a small group of economic elites. There are, of course, other reasons that Iran is not democratic (including possible cultural reasons), but Karl offers one important reason that we would not expect political change there anytime soon. In this way, research on the natural resource curse makes a compelling case that even if education and media become more prevalent, states relying primarily on a single natural resource are unlikely to become democratic.

In reviewing existing approaches, Michael Ross (1999) points out several distinct theories for why oil-rich states have difficulty establishing democratic systems of government and notes that previous research on the natural resource curse does not offer insight into which theory is correct (factors that some theories, such as Karl’s, do not

acknowledge). One explanation is called the *rentier effect*, which proposes that governments give special privileges to the oil industry to avoid being accountable to the general population in order to stay in power. Another explanation is the *repression effect*, in which governments use oil revenue to develop their internal security, thereby repressing political participation and popular demands for public goods. Finally, there is a *modernization effect*, which harkens back to traditional modernization theory, arguing that oil wealth does not encourage the social and cultural shifts thought to underpin successful democracies (Ross, 2001).

Relying on statistical analysis of 113 countries from 1971 to 1997, Ross finds evidence for all three of these theories of why oil-rich states might not promote democracy. Notably, he emphasizes that the effects of the natural resources curse are not limited to the Middle East or oil-based economies. He finds not only that oil is bad for democracy but that the presence of large amounts of any mineral resource impedes democratic societies, as can be seen in countries as diverse as Angola, Chile, Cambodia, Congo, and Peru.

Research on the natural resource curse suggests that we must consider the type of economic development occurring and its specific consequences in order to understand the effect on democracy. Another explanation for the existence of wealthy dictatorships—which is consistent with Ross’s argument—is that wealth makes any political system less likely to change, whether that system is democratic or not. This means that wealthy dictatorships are likely to remain dictatorships and wealthy democracies are likely to stay democratic. This is the basic argument of *survival theory*.

Survival Theory

Modernization theory, having faced some significant objections, was brought back to life, at least partially, by proponents of survival theory. According to Przeworski and Limongi (1997), the problem with traditional modernization theory is that it observes that democracies tend to be wealthier than nondemocracies but cannot account for two conflicting explanations of this observation. On one hand, as countries become wealthier, they may become more likely to establish democratic governments. On the other hand, democracies may emerge regardless of the level of wealth but are more likely to last when they are wealthier. We would observe the same outcome—that democracies tend to be wealthier—in either case. Based on this outcome, we cannot simply assume, however, that wealth is the cause of democracy, because wealth may merely be the cause of democratic survival, not its emergence.

Przeworski and Limongi (1997) argue that rather than economic development’s being responsible for the emergence of democracy, wealth makes all political systems, including democracies and dictatorships, more likely to endure. If this is true, then we should expect democracies to arise irrespective of their level of wealth but to be less

likely to last when they are poor. Likewise, poor dictatorships are likely to face instability although not necessarily a transition to democracy. By this account, democracy “survives if a country is ‘modern,’ but it is not a product of ‘modernization’” (p. 159). Przeworski and Limongi conducted statistical analysis on an impressive data set of 135 countries from approximately 1950 to 1990, tracking when a country became a democracy or dictatorship, how long it lasted, and the levels of development and growth rates. Their findings reveal that as dictatorships become wealthier, a transition to democracy is more likely, but only up to a point. Once dictatorships reach very high levels of economic development, they are remarkably stable, and democracy is unlikely to emerge. Furthermore, Przeworski and Limongi found—contrary to what Lipset and Huntington would have predicted—that rapid growth is not destabilizing for democracies. Relatively poor countries with some economic growth are more stable than rich democracies that experience decline.

In summary, survival theory does not take a strong stance on the reasons countries become democratic in the first place, but rather it emphasizes that democracies arise regardless of the level of development (for whatever reason) but are more likely to be sustained when the level of wealth is high.

Summary of Lessons From Modernization Theory

Despite extensive study on the connection between politics and economics, scholars are still divided about the relationship between political development and modernization. One major reason for this uncertainty is that researchers often struggle with a dearth of data that is comparable across countries. Some measures of economic and political characteristics are available only for particular regions or periods, thereby limiting the scope of a project, or researchers may be interested only in specific regions. Przeworski and Limongi (1993) list some of the major studies, conducted from the early 1960s through the 1990s, that tried to determine the connection between democracy and economic development. These studies include as few as 10 underdeveloped countries or as many as 100 countries. Some studies are limited to a single region, such as Latin America, while others are more cross-national, including several regions. The time over which these countries were studied varies as well. The inevitable result of drawing conclusions based on different countries during varying time periods is a myriad of results, some of which are contradictory.

There are, however, a couple of lessons that we can draw from the wide array of research examining political development and modernization. First, economic growth is probably not necessary for the emergence of democracy. There is compelling evidence that democracy emerges at all levels of development. At a minimum, we can safely say that the economy alone does not predict democratization.

Second, economic growth may aid in the survival of democratic governments, but it is likely not sufficient. Other factors matter too, as evidenced by the so-called natural resource curse. For instance, excessive inequality may hamper citizens' ability to participate and may encourage the development of institutions dominated by a few powerful elite rather than an active civil society. Furthermore, Moore's (1966) central thesis is that only certain paths of modernization result in democratic systems. Robert Bates (1991) similarly lays out different paths of economic modernization, some of which explain the developments of historical Europe and others, the experiences of formerly socialist countries.

These lessons are hedged in probabilistic terms because much research remains to be done to definitively establish the political consequences of economic changes. One of the most important lessons that we can take from this research is that when explaining the connection between politics and the economy, we must fully consider the nature and pace of economic development and its specific consequences. Specifically, we must consider which groups are empowered or hurt by growth or decline, the incentives of the state and politicians under such circumstances, and the behavior of citizens. Only by addressing these intermediary links can we fully understand the connection between modernization and political development.

Policy Implications for Promoting Democratization

Research on political development and modernization has often been explicitly motivated by a desire to shape foreign policy, revealing a clear connection between academic theories and the policy realm. When writing *The Civic Culture* in the early 1960s, Almond and Verba were responding to the spread of Communism around the world and attempting to identify the necessary prerequisites for democratic systems. They even expressed concern that some Western European nations would fail to find a stable form of democracy. Likewise, Rostow's *The Stages of Economic Growth* was a guide for the U.S. Agency for International Development, which sought to formulate a policy to prevent the spread of Communism to South Vietnam and Indonesia. More recently, Huntington's *The Third Wave: Democratization in the Late Twentieth Century* (1991) includes sections devoted to advice for politicians seeking to establish democracies in their own countries.

Depending on which theories we accept, prescriptions for foreign policy and democracy promotion around the world change drastically. Modernization theory in its original formulation suggests that those interested in promoting democratization around the world should encourage countries to pursue policies of economic development as a minimum prerequisite for a democratic society. However, research suggests that the ways in which governments promote economic development will influence whether the

net effect actually promotes democratization. As suggested by survey research, economic growth alone may not be sufficient to produce democratic systems of government.

In *The Logic of Political Survival*, Bruce Bueno de Mesquita, Alastair Smith, Randolph Siverson, and James Morrow (2003) offer a different way of conceptualizing types of political regimes, and their approach has important implications for foreign policy. According to traditional definitions of democracy, countries are democratic if they hold free and fair elections and meet other important procedural requirements, including a free press and civil liberties. However, Bueno de Mesquita et al. put forth the *selectorate theory* and argue that the crucial characteristic for understanding how a government will behave lies in knowing to whom leaders are accountable. The *selectorate* are all those who have a say in the selection of the leaders (and that selection may or may not be through voting in democratic elections). The *winning coalition* is the group of individuals whose support is necessary for a leader to stay in power. For instance, in a democracy, the selectorate is very large, encompassing all possible voters, and the winning coalition consists of the majority necessary to win an election. In contrast, in a military dictatorship, the selectorate is likely to be a small group of officers, and the winning coalition is the minimum number of those officers necessary for the dictator to maintain power.

Distinctions between the size of the selectorate and the winning coalition are significant because they have important implications for the types of goods leaders will provide to their citizens. When there is a large winning coalition and a large selectorate, as in democracies, leaders have an incentive to provide public goods in order to maintain the support of the large proportion of the population needed to win elections. Conversely, when there is a small selectorate and a small winning coalition, it is more efficient for leaders to provide private goods to the small group on whom they rely for power.

This logic of political survival suggests that foreign aid—even if it is designed to promote economic growth—may have the reverse effect by providing dictators with resources that can be diverted to a small winning coalition rather than used to promote development. Foreign aid, therefore, may not be favorable to the type of economic development that promotes the establishment of democracy if it goes to a dictator who can pay off supporters. This is why it would not have been advisable to give foreign aid directly to Saddam Hussein, the former dictator of Iraq, who could have used it to his own ends and specifically to prop up support for his regime among his small winning coalition. Doing so would have made him and his few supporters more powerful without encouraging the growth of a middle class or the broader education of the population, which we hope provides the basis for a democratic society. Rather than giving aid directly to Hussein, the United Nations attempted to set up an “oil for food” program by which oil revenues were traded for food provided directly to the Iraqi people.⁸

Existing research suggests that although modernization may lead to democratization, it is likely a condition that is neither necessary nor sufficient. Selectorate theory suggests that we should seriously consider the incentives of political leaders to provide public goods and create free democratic systems. Other prominent works, such as that of Daron Acemoglu and James Robinson (2006), *Economic Origins of Dictatorship and Democracy*, suggest that the success of democracy depends on elites' not having incentives to oppose it. This research indicates that in developing foreign policies intended to promote democratization, we should not limit ourselves to promoting economic development and broader education among the population but should explicitly consider the motivations driving political leaders.

Future Directions

There are many remaining questions about the connection between political development and modernization. There are three areas at the forefront of research about the link between politics and economics—understanding the shared foundations of development and freedom, expanding our knowledge about the wide variety of nondemocratic systems, and better understanding the preferences of groups in reaction to socioeconomic changes.

First, scholars have recently suggested that there are shared foundations for both economic growth and democracy. For instance, James Robinson (2006) suggests that there are underlying causes—such as secure property rights and the rule of law—that explain why countries are economically successful and more democratic. The result is that we observe that democracies are typically wealthy, not because wealth causes democratization, but because the same factors make both more likely. In a similar vein, Amartya Sen (1999) advances the idea of shared underlying causes of democracy and economic growth by making the case that development should be conceptualized as removing “unfreedoms” from society. In writing that “development is indeed a momentous engagement with freedom’s possibilities,” he expresses the idea that individual freedom is both the end and the means of development (p. 298). Only by ensuring a social commitment to the protection of individual freedom can societies progress. This is a radically different and integrated conception of political and economic development, one that is largely responsible for Sen’s Nobel Prize in economics.

A second area of future research examines the tremendous variation among nondemocratic systems. This chapter has focused on the determinants of democracy without considering the many different types of nondemocracies. However, research has suggested that nondemocracies vary in politically significant ways. For instance, Juan Linz and Alfred Stepan (1996) trace differences in the

democratic transitions of nondemocratic governments, which they classify as *authoritarian*, *totalitarian*, *posttotalitarian*, and *sultanistic*. However, many questions remain, including why military dictatorships arise in some countries but rigged-election systems emerge in others.

Selectorate theory offers a starting point for understanding the differences among nondemocratic systems. Bueno de Mesquita et al. (2003) note that some nondemocracies are more successful than others at achieving economic development and providing citizens with public goods such as roads and schools. The reason is that the loyalty of a leader’s winning coalition—the people on whom the leader relies for power—varies markedly in different nondemocratic systems. For instance, in dictatorships with rigged elections, it would be difficult for the small group of supporters (who help the leader cheat in the election) to defect and bring another ruler to power because there are many people available who can help rig an election. The result is that the leader does not need to spend a lot to buy the support of the winning coalition because it provides a very loyal base. Because the leader does not need to spend a lot to maintain support, he does not need to be overly concerned with producing national wealth through economic development. In contrast, in a military dictatorship in which the winning coalition (a handful of military officials) can successfully and relatively easily transfer their support to another dictator, the leader needs a lot of resources to continue to buy support. As a result, the leader in a military dictatorship has more incentive to promote economic development in order to obtain the wealth necessary to pay off a more fickle group of supporters. This research provides a starting point for understanding variation among nondemocratic systems, but much work remains to be done on the origins of the many different types of nondemocratic systems and other ways in which their political systems vary.

A final area of future research is inquiry into how groups such as private businesses and citizens react to socioeconomic changes. Groups should not simply be lumped together as a whole in studying the connection between politics and economics. A notable reason is evident in research on the welfare state. Businesses are often thought to oppose national welfare programs, whereas citizens are assumed to support them. However, Isabela Mares (2003) notes that businesses have not always opposed the introduction of welfare state programs, as is often assumed, but that their preferences depend on the kind of business. Business support for national welfare programs hinged on whether businesses would have control over the national system and whether they perceived benefits in the reduction of risk (due to illness and old age in the workforce) that could be beneficially distributed across all businesses.

Likewise, there is evidence that citizens are not always opposed to retrenchments in the welfare state. In *Democracy and the Market*, Przeworski (1991) frames the

dilemma of market-oriented reforms in Latin American and post-Communist countries of Eastern Europe as being one of forward-looking politicians attempting to overcome the opposition of citizens hesitant to accept the uncertain costs of transitions. However, evidence suggests that there are instances in which citizens can be convinced of the benefits of market-oriented reforms.⁹ Such research suggests an interesting direction for a better understanding of the ways economic actors play a role in political developments.

Advancing our understanding of the connection between political development and modernization requires us to be more nuanced in our understanding of what constitutes development, the different types of nondemocratic systems, and the diversity of group responses to economic changes.

Conclusion

This chapter has reviewed the major developments in research on political development and modernization. Global problems persist in highlighting the importance of studying why and how economic and political systems should be designed. According to Freedom House, as of 2008, a little more than half of all people in the world (about 3.6 billion) lived in political systems that are not free. Although the World Bank notes some progress on poverty alleviation—from 1981 to 2005, rates of extreme poverty have declined from more than half to about a quarter of the global population—this means that about 1.7 billion people continue to live without their basic material needs being met.¹⁰ Addressing why these problems arise and persist is of paramount concern in achieving freedom from hunger and repression—two basic, yet still elusive goals, for most of the world's population.

Notes

1. Aristotle. *Politics*. Book VI, Chapters 2–3. In R. McKeon (Ed.), *The Basic Works of Aristotle* (pp. 1265–1268). New York: Random House.

2. Alexis de Tocqueville. (1835–1840). *Democracy in America* (H. Reeve, Trans.). London: Saunders & Otley.

3. Francis Fukuyama. (2006). Foreword. In S. Huntington, *Political Order in Changing Societies* (p. xiii). New Haven, CT: Yale University Press.

4. Jorge I. Domínguez (2001) explains the different interpretations of Huntington's work, including the ways in which he challenged modernization theory while simultaneously using an approach relying on factors highlighted by modernization theory. See Domínguez, J. I. (2001). Samuel Huntington and the Latin American state. In M. A. Centeno & F. López Alvez (Eds.), *The Other Mirror: Grand Theory Through the Lens of Latin America* (pp. 219–239). Princeton, NJ: Princeton University Press.

5. Valenzuela and Valenzuela (1978) compare the assumptions, methods, and conclusions reached by modernization and dependency theory in the context of Latin America. Writing in 1978, they concluded that modernization theory is more parsimonious but makes arbitrary distinctions about which phenomena are traditional, thereby explaining a lack of development. In their view, dependency theory drew on a broader set of evidence but needed to be more precise in laying out its concepts and causal explanations.

6. In explaining proper research design, Barbara Geddes offers a very useful summary of the rise and fall of dependency theory and an explanation of why the theory lasted for so long despite compelling cross-national evidence. See pp. 6–17 in Geddes, B. (2003). *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.

7. See "Iran Expands Role in Media, via Satellite and in English." *The New York Times*, July 3, 2007. Available at <http://www.nytimes.com/2007/07/03/world/middleeast/03iran.html?scp=3&sq=Iran%20media&st=cse>

8. The former Secretary General of the United Nations, Kofi Annan, defended the program as one of the most successful and unusual development programs in that it relied on the resources of the country itself to provide aid. His full remarks, given in 2003 when the program was ended, are available at <http://www.un.org/Depts/oip/background/latest/sgstatement031119.html>

9. One concrete example can be seen in pension reform measures in Poland. Pension privatization measures radically shifted the burden of providing for old-age retirement from the state to citizens. Nonetheless, survey research shows that Polish citizens were at least partially supportive of implementing the new system, which had been justified in terms of being fairer by creating a link between contributions and benefit. See Chłóń, A. (2000). *Pension Reform and Public Information*. Social Protection Discussion Paper No. 0019. Washington, DC: World Bank.

10. Freedom House provides data and analysis tracking freedom around the world, available at www.freedomhouse.org. The World Bank and United Nations provide comprehensive data on the many facets of poverty. See the United Nations' Human Development Reports at <http://hdr.undp.org/en/>. The World Bank's World Development Indicators are available through <http://www.worldbank.org>.

References and Further Readings

- Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge, UK: Cambridge University Press.
- Almond, G., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.
- Bates, R. (1991). The economics of transitions to democracy. *PS: Political Science & Politics*, 24(1), 24–27.
- Boix, C. (2003). *Democracy and redistribution*. Cambridge, UK: Cambridge University Press.
- Bueno de Mesquita, B., Smith, A., Siverson, R. M., & Morrow, J. D. (2003). *The logic of political survival*. Cambridge: MIT Press.

- Deutsch, K. (1953). *Nationalism and social communication: An inquiry into the foundations of nationality*. New York: Wiley.
- Deutsch, K. (1961). Social mobilization and political development. *American Political Science Review*, 55(3), 493-514.
- Frank, A. G. (1970). The development of underdevelopment. In R. I. Rhodes (Ed.), *Imperialism and underdevelopment: A reader* (pp. 4-17). New York: Monthly Review Press.
- Huntington, S. (1968). *Political order in changing societies*. New Haven, CT: Yale University Press.
- Huntington, S. (1991). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Ingelhart, R. (1990). *Culture shift in advanced industrial society*. Princeton, NJ: Princeton University Press.
- Karl, T. L. (1997). *The paradox of plenty: Oil booms and petro states*. Berkeley: University of California Press.
- Lerner, D. (2000). The passing of the traditional society. In J. T. Roberts & A. Hite (Eds.), *From modernization to globalization: Perspectives on development and social change* (pp. 119-133). Malden, MA: Blackwell. (Original work published 1958)
- Linz, J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and post Communist Europe*. Baltimore: Johns Hopkins University Press.
- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53, 69-105.
- Lipset, S. M. (1960). Economic development and democracy. In S. M. Lipset (Ed.), *Political man: The social bases of politics* (pp. 27-63). New York: Anchor Books.
- Mares, I. (2003). *The politics of social risk: Business and welfare state development*. Cambridge, UK: Cambridge University Press.
- Moore, B. (1966). *Social origins of dictatorship and democracy: Lord and peasant in the making of the modern world*. Boston: Beacon Press.
- Parsons, T. (1971). *The system of modern societies*. Englewood, NJ: Prentice Hall.
- Przeworski, A. (1991). *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. Cambridge, UK: Cambridge University Press.
- Przeworski, A., & Limongi, F. (1993). Political regimes and economic growth. *Journal of Economic Perspectives*, 7, 51-69.
- Przeworski, A., & Limongi, F. (1997). Modernization: Theories and facts. *World Politics*, 49, 155-183.
- Robinson, J. (2006). Economic development and democracy. *Annual Review of Political Science*, 9, 503-527.
- Ross, M. L. (1999). The political economy of the resource curse. *World Politics*, 51, 297-322.
- Ross, M. L. (2001). Does oil hinder democracy? *World Politics*, 53(3), 325-361.
- Rostow, W. W. (1959). The stages of economic growth. *Economic History Review* (new series), 12(1), 1-16.
- Rueschemeyer, D., Stephens, E. H., & Stephens, J. D. (1992). *Capitalist development and democracy*. Chicago: University of Chicago Press.
- Sen, A. (1999). *Development as freedom*. New York: Anchor Books.
- Shafer, M. (1994). *Winners and losers: How sectors shape the developmental prospects of states*. Ithaca, NY: Cornell University Press.
- Valenzuela, S. J., & Valenzuela, A. (1978). Modernization and dependency: Alternative perspectives in the study of Latin American underdevelopment. *Comparative Politics*, 10, 535-557.

11

STATISM

ANNA BATTÀ

University of North Texas

Prior to World War II, scholarship in comparative politics focused mainly on the study of institutions in western European countries, and formal theories examined the workings of the state in detail. At that time, the state was viewed as an autonomous entity with a large degree of power on its own. Society and its impact on the state seemed somewhat unimportant. The behavioral revolution of the 1960s and 1970s, on the other hand, altered the focus of political inquiry completely. With the rise of the *society centeredness* approach, individual and group behavior assumed primacy of analysis, and political scientists retreated to a large extent from the study of the state. Indeed, scholarship of this time portrayed the state as a black box without having its own interests and being merely an arena where behavioral games of various groups and individuals were acted out. Possibly the biggest contribution of recent statist theories is the idea of “bringing the state back in” (as Skocpol, 1985, expressed it) to the attention of scientific inquiry. Theda Skocpol’s works (1979, 1985) began this movement, which was later extended by emphasizing the state’s security function (Ayoob, 1992; Tilly, 1985) and diversity of structure in the various forms of autocracies, such as corporatist (Malloy, 1977), bureaucratic (O’Donnell, 1978), neopatrimonial (Bratton & van de Walle, 1994), and totalitarian regimes (Bova, 1991; Bunce, 1999). A focus on weak states (Migdal, 1988) and the politician’s dilemma in developing countries (Geddes, 1994) completed the general scholarship

on providing a balanced view between individual and group behavior, as well as the workings of the state.

In general, statist theories define what is understood by the institution of the state and analyze the relationship between state and society. In addition, authors within this subset of literature provide explanation about the formation and development of states. Despite its frequent usage, the concept of the state has been ambiguous in the social sciences and has been defined arbitrarily in the past (Sabine, 1934, cited in Almond, 1988). Perhaps because of its strong-fisted or authoritarian connotation, as in “*l’état c’est moi*” by Louis XIV, and the emergence of the norm of state sovereignty after the Treaty of Westphalia, political scientists have preferred to call the state anything but that (Watkins, 1968). As a result of far-reaching political mobilization in the West during the 19th and 20th centuries, the term *state* was replaced by *government* and later on by *political system* (Almond, 1988). State sovereignty and the emergence of new political institutions, such as political parties, pressure groups, and the mass media, facilitated a change that called for a broader term for referring to the general entity of the state.

The notion of the state, however, remained important in the Marxist and neo-Marxist interpretation of politics. Scholars working within this paradigm view the state as the “coercive instrument of the ruling class,” defined in terms of the ownership and control of the means of production (Miliband, 1969, p. 5). Essentially, under the

capitalist system, the members of the ruling class control the political system and determine the state's policies and actions (Miliband, 1969). Because of the above dynamics, political equality, save in formal terms, is impossible in the conditions of advanced capitalism, Marxists argue. Therefore, the state as a powerful entity plays an important role in supporting the capitalist class and in perpetuating the status quo.

Eric Nordlinger (1981) aims at avoiding anthropomorphizing the state and settles for a definition that stresses the policy-making function of individuals rather than institutional arrangements. Based on his understanding, the definition of the state must focus on individuals because it is only individuals who are able to create public policy.

State Organization and Domination

One of the earliest contributors of statist theories, Max Weber, in a speech given at Munich University in 1918, lays out what became known as the definition of the modern state. According to Weber, the state "cannot be defined in terms of its ends; . . . ultimately, one can define the modern state sociologically only in terms of the specific means peculiar to it, as to every political association, namely, the use of physical force" (Gerth & Mills, 1946, p. 77). Prior to the establishment of the modern state, the use of physical violence by actors other than the state was considered normal because territories in Europe changed hands frequently. However, with the rise of the modern state, the use of physical force was gradually restricted to the state. In other words, the state assumed "the monopoly of the legitimate use of physical force within [its] territory" (Gerth & Mills, 1946, p. 78).

State domination, however, rests on additional factors, such as traditional, charismatic, and legal aspects of legitimation, Weber argues. Rules and historically recognized habits that became part of national consciousness are believed to have special power that can in itself bring about compliance by the populace. This argument is not unlike Montesquieu's (1989) in *The Spirit of the Laws*, in which he states that obedience to a power is sustained not necessarily because of a rule's existence but because of a long-standing history and adherence to that particular rule. Further, citizens of a country may also follow the leader's decision if he or she is gifted with extraordinary charisma or prophetic ability. Finally, popular domination may also take place via the legal system and rationally established rules. Perpetuating either form of this domination is a function of controlling the material goods the state is equipped with and acquiring additional material from tax revenues collected by the state. Organized domination requires administrative staff, which is bound to the power holder by either material reward or social honor or both (Gerth & Mills, 1946).

In agreement with Weber, subsequent scholars of statist theory argue that the most important function for which the

state organizes is its security function (Ayoob, 1992; Tilly, 1985). With the famous phrase "War makes the state," Charles Tilly (1990) argues that states come about for the main reason of defending themselves from other states. Based on this understanding, the government's main function is to monopolize violence, which inherently functions on economies of scale. In other words, within a particular country, the state is the sole possessor of violence for the primary reason of coercion, and since this function encompasses the entire territory, monopolized violence is carried out on a large scale. According to statist theories, the state is engaged in legitimate organized crime for the protection of individual citizens, who have no choice but to accept this public good (Tilly, 1985).

According to Tilly (1985), agents of states have four types of activities in terms of organized violence, which are war making, state making, protection, and extraction. This assessment is an extension of Lane's (1958) analysis of protection. The purpose of *war making* is to eliminate or to neutralize the state's "own rivals outside the territories in which they [i.e., the state] have clear and continuous priority as wielders of force" (Tilly, 1985, p. 181), whereas *state making* is aimed at "eliminating or neutralizing their rivals inside those territories" (p. 181). The *protection* activity intends to eliminate or neutralize the enemies of the state's clients; and finally, *extraction* capability refers to "acquiring the means of carrying out the first three activities—war making, state making, and protection" (p. 181).

Bringing the State Back In

During much of the behavioral revolution in the social sciences, analysis mainly focused on individual and group behavior, and the society-centeredness approach dominated major discourses. This approach denied the autonomy of the state and viewed the state as a place in which various actors seek to dominate politics. Beginning with Skocpol, however, the perception of the state began to change. According to Skocpol, the state is very much an autonomous unit, and it has interests of its own as an organic entity (Skocpol, 1979, 1985). Specifically, the state has the ability to generate situations in which revolutions can occur, and rebels are able to exploit such situations. Skocpol questions the core idea of the society-centeredness approach, as well as the Marxist view that the state's function is nothing else but serving the interest of the dominant economic class. According to Skocpol, the state is autonomous from the dominant economic class, with interests of its own that are different from the dominant class (Mason, 2004). Essentially, the state's function is not merely serving the interests of the ruling capitalist class.

Authors besides Skocpol moved away from the society-centered viewpoint (Eckstein, 1960; Evans, Rueschemeyer, & Skocpol, 1985; Krasner, 1984; Truman, 1951).

Stephen Krasner (1984) argues that if the state and public institutions are merely seen as referees of politics, this would negate the idea that the state is able to manipulate and change its own society. The major function of public officials is perhaps more than only making sure that the game is played fairly; the state is therefore an autonomous actor in the political system. Further, given that executives, congresspeople, and judges each have authority and discretion in their decision making, governmental institutions do have autonomous decision-making capacity (Truman, cited in Almond, 1988).

Yet other scholars suggest that the state has even more autonomy than previously stated. In addition, they argue that the pluralist paradigm is not one-sided; a two-way relationship exists between the state and society (Almond, 1988; Eckstein, 1960). According to Gabriel Almond, both the society-centeredness viewpoint and the bringing-the-state-back-in group are missing the point. Within a pluralist system and widespread representation, society has the ability to exert a large degree of influence on the state, but the state also has the ability to influence society. Austin Ranney (1966), for example, gives the example of the president who is the chief legislator and implies that this function is rarely a passive one. Further, the political executive must deal with a variety of interest groups and has to be able to resist the pressures stemming from them (Carter & Hertz, 1972). Finally, bureaucrats are responsible for the majority of policy initiatives and take part in the design of policies, as well as influencing the agenda of decisions (Putnam, 1973).

Nordlinger (1981; Nordlinger, Lowi, & Fabbrini, 1988) seems to provide somewhat of a moderate middle ground between the two views, although his stance is not entirely clear. He states that the ability of public officials to translate their own preferences into public policy is fairly limited if those preferences deviate largely from the preferences of society. Almond perhaps has gone too far by attributing too much autonomy to state officials, he argues. Officials' ability to turn nominal power into decisional power is severely restricted. Nordlinger dismisses the idea that the state is an all too powerful, autonomous entity able to assert its own policy preferences. Instead, he argues that the state is a malleable unit that is indeed "in the hand of the most powerful individuals" (Nordlinger, 1981, p. 21). Only when preferences of state and society converge will the state follow the preferences asserted by society. However, there is a two-way relationship between these two actors, and in agreement with Almond, Nordlinger states that the society-centered models ignore this two-way relationship.

David Easton (1981) states that the return to the state movement arises out of a contemporary revival of Marxism. Yet those working within this movement differentiate themselves from Marxists because they clearly articulate that the state is more than merely subservient to the ruling capitalist class (Skocpol, 1979, 1985). Based on the Marxist conception of the state, owners and controllers

determine policies, and the state is never an autonomous entity and does not have the ability to act against the interest of the dominant capitalist class. Neo-Marxists, on the other hand, argue that the state is relatively autonomous and temporarily acts against the interests of the dominant class to maintain capitalism, which is the long-term interest of the dominant class. In essence, the state may act against the short-term interests of the ruling class in order to save capitalism.

State Building and Structure

In addition to defining what is included in the idea of the state and clarifying the relationship between state and society, statist theories assess the difficulties state formation entails in both industrialized countries and developing nations. The formation of the state in the West was characterized by a long, gradual, but also violent process during which the frequent exchange of territories eventually came to a halt and borders solidified. With the signature of the Treaty of Westphalia, the international system accepted an entirely new arrangement, based on which the building blocks of the system were nation-states equipped with full sovereignty and the promise of noninterference by other equally sovereign states. What exactly is meant by *nation state*, however, has not been clear since. On one hand, *statehood* refers to the political side of a country, including especially the institutional aspect of governance, whereas a *nation* is defined as encompassing a relatively uniform culture and ethnicity, as well as identity within a country, in addition to the political aspect. Furthermore, based on Article 3 of the Montevideo Convention, which restated already existing international law, a country is considered a nation-state if it has (a) a permanent population, (b) a defined territory, (c) a functioning government that has control within the borders of the country, and (d) a capacity to enter into relations with other states. Even though many countries, such as many weak states in Africa, for a certain period in their histories clearly do not meet some of these criteria, these standards have been typically the baseline of evaluation for nation-state status.

In an increasingly interconnected world in which global trade and economic activity are widespread and in which increased transgovernmental organizational activity prevails, it is important to evaluate how recognition of individual states by the international community factors into the definition and assessment of nation-state status. At present, an internationally accepted treaty that would achieve this aim does not exist. The above-mentioned Montevideo Convention does not incorporate this element and resorts instead to the "declarative theory of statehood," omitting the recognition specification (Hillier, 1998). Some nations subscribe to the "constitutive theory of statehood" (Lauterpacht, 1958), which does intend to add international recognition; however, this theory has been criticized on the

grounds that it is incomplete and does not deal with instances in which states are not recognized by others.

Despite the above difficulties with respect to clarifying the state as the unit of the international system, developing countries were required to build nation-states based on preexisting conditions that barely resembled those of the Western state-building process. Mohammed Ayoob points out important differences between third world countries and established Western democracies in terms of state structure and organization during state formation. Third world countries not only are weak states, which means that they have a much more difficult time penetrating society, but they also face civil wars within their borders with a much higher frequency than industrialized nations had to during the time of their state formation (Ayoob, 1992; Migdal, 1988). Resource extraction is also problematic for newly emerging states because of widespread poverty, and institution building is often stalled by high levels of corruption and strong ethnic differences.

State building is shaped and determined in the context of the international environment. Statist theories point out that third world countries are faced with the task of establishing a nation-state because it is the only acceptable unit of political interaction on the broader world stage. The international system is set up with the state as a building block, and whether some countries' communal identities or organizational structures match the system or not, they have to follow the prescribed scheme nevertheless. It is argued that nationalist identity in countries following colonization was frequently forged, and the sense of community did not follow already existing ethnic lines (Ayoob, 1992).

Again, the development of the nation-state in Europe differed drastically from the formation of the state in other areas of the world, especially the third world. State development in Europe took centuries and was accompanied by bloody civil wars and ethnic as well as religious disagreements. Perhaps it is too much to expect third world countries on the path of nationhood not only provide security during the transition period but also offer adequate welfare and enable participation at all levels of government, each of these elements being a difficult task on its own. European countries at the time of their state formation faced merely the provision of security. Developing countries, on the other hand, are expected to achieve a much higher level of development at a considerably faster pace and at the same time fulfill the many demands that a globalized, modern, and complex world places on them.

Types of Authoritarian States

Another contribution of statist theories is that, in addition to the standard dichotomy of democratic versus authoritarian regimes, they differentiate among various subtypes of authoritarian states and assess the problems

associated with such arrangements, as well as recommending policy for dealing with those problems. Five major types of authoritarian state systems can be distinguished: corporatist, bureaucratic-authoritarian, neopatrimonial or sultanistic, totalitarian, and posttotalitarian (Linz, 2000). The prospects of democracy differ according to which state system operates in any given country.

The *corporatist* state is typically a one-party-dominant system in which only a limited number of component units or sectors exist, such as labor unions, peasants, merchants, and industry, and no competition takes place among specific units (Malloy, 1977). Essentially, the state ultimately decides which units will be allowed to participate in the system, and the state strictly controls leadership selection of the individual units. Membership in a corporatist setting is compulsory, and if a person, by occupation, is a member of a certain category, he or she is automatically a member of the organized corporate group. Organization outside of the boundaries of corporate groups is not possible. In addition, the structure of individual compartment units, such as labor, industry, peasants, merchants, and so forth, is hierarchical, and only the leadership has direct interaction with the agents of the state.

Furthermore, the corporatist system is characterized by a strong element of bureaucratic planning, and although elections do take place, it is clear which party will be the winner of those elections. Examples of corporatist authoritarian states include Mexico under the Institutional Revolutionary Party (PRI) and the South Asian Tigers, such as Singapore, South Korea, Taiwan, and Hong Kong. Corporatism in these states functions as a way of accommodating and accepting the process of modernization and development. Stresses that naturally come from modernization are believed to be relieved if the state manages society adequately. Conflict is not inevitable; however, if the state is able to possess the authority and the skills to manage the process of modernization, conflict can be minimized. Democracy has the highest probability of springing up in the corporatist system, given that individual compartment units are already organized and can serve as the basis of civil society and democracy if they are able to cast off their ties to the dominant party (Malloy, 1977).

Within corporatist settings, the state typically provides benefits, such as steady jobs, to those who are part of the various sectors. In case the state fails to sustain the provision of such benefits, the system evolves into a *bureaucratic authoritarian* system, in which elections are suspended entirely (O'Donnell, 1978). Bureaucratic-authoritarian regimes evolve when groups that had been previously actively involved in policy making are deactivated, which is often achieved via military coup. Transition to democracy from a highly technocratic state, such as the bureaucratic system, can take place after the old corporatist order is restored.

In a *neopatrimonial* state, the right to rule is prescribed to a person and not to an office, and benefits are given in

exchange for political loyalty (Bratton & van de Walle, 1994). Political office in such a system is determined by loyalty and is not based on merit. In addition, bureaucratic power is either weak or nonexistent. Neopatrimonial states typically do not display the formal governing coalitions between state and social interests or the collective bargaining over core policy issues that are found in corporatist states. In neopatrimonial states, the chief executive maintains authority through personal patronage rather than ideology or law.

Observable differences exist between the corporatist and the neopatrimonial system. Under neopatrimonialism, for instance, the state has undermined capitalist forms of accumulation, whereas in a corporatist state, capitalism is allowed to function. In other words, under corporatism, the state respects individual property rights and observes the rules of the capitalist economy, but in the neopatrimonial state, private ownership of assets is not the norm, and state ownership of such assets prevails.

Both corporatist and neopatrimonial regimes are built on personalistic ties and typically experience high levels of corruption. The *patron–client* relationship is pointed out by John Duncan Powell (1970), who argues that environmental scarcity and the constant threat of drought or other natural disasters make peasants vulnerable economically, and therefore, to secure protection, peasants turn to the patron on a regular basis. There is a degree of power asymmetry between the superior (patron) and the subordinate (client). The relationship between these two actors depends on reciprocity in the exchange of goods and services. On one hand, the patron provides security for those engaged in subsistence farming, and the client reciprocates with political support for the patron. Patron–client relations are strong and typically enduring in traditional villages, where face-to-face contacts prevail. In a more integrated village, the relationship is characterized by differentiation and specialization. This type of contact is highly personalistic, largely with informal agreements. Essentially, a public entity that would function as an enforcement authority does not exist. There are no sanctions of compliance, and transactions take place within the realm of private accountability. This form of association differs largely from the relationship between citizen and representative, in which public accountability and the rule of law prevail. Peasant unions under the patron–client system either do not form or are organized from above, such as in the corporatist state (Scott, 1972).

The final two subsets of authoritarian systems are *totalitarian* and *posttotalitarian* regimes. Totalitarian systems are characterized by state ownership and control of all production assets and the complete absence of market forces (Bova, 1991; Bunce, 1999). These types of systems employ centralized planning, and one party dominates the political agenda. The regime has eliminated all preexisting political, social, and economic pluralism and has a unified, articulated, and guiding utopian ideology.

Leadership in a totalitarian setting functions with a high degree of uncertainty.

Posttotalitarian regimes are similar to totalitarian systems, with the major difference that the central ideology fades over time. Under posttotalitarianism, the state still has one party and one party only; however, some constraints on the leader are established gradually, and there is a more noticeable degree of social pluralism. Leaders of such regimes rule less by charisma and more on the basis of technical competence. Bureaucratic and technocratic leadership is more the norm than highly ideological and charismatic control. In some cases, elements of democratic opposition can be present in civil society and could lead to democracy. These elements, however, are not allowed in a totalitarian state. Furthermore, the black-market economy is tolerated as an alternative to the inefficiencies of a state-managed economy. It is possible for a partial market economy and a planned state economy to coexist within a posttotalitarian state (Linz & Stepan, 1996).

Transition from a totalitarian regime to a posttotalitarian state is visible once the overarching ideology declines and state legitimacy wanes as a result. Increasingly a growing disjunction is present between the official ideology and reality (Linz & Stepan, 1996). Citizens and elites are less committed to state ideology, and growing criticism of the regime comes from society. Although there is a clear decline in ideology, the leadership still observes a coherent set of ideas, unlike in other authoritarian regimes. Transition to democracy from posttotalitarianism is complicated by the lack of a preexisting civil society and the absence of political parties and democratic leadership. Obtaining a democratic political culture in this case is a difficult task.

The Politicians' Dilemma and Weak States Syndrome

Early statist theories focused on the institutions of the state whereas works written during the behavioral revolution purposely played up the role of the grassroots level. Very often the literature leaves out a balanced description of the relationship between state and society. This balance, however, is better accomplished in scholarship dealing with the third world, such as scholarship exploring the dilemma politicians face when trying to deal with the limitations of a patron who stands as a filter between state and society (Geddes, 1994) and the weak states syndrome (Migdal, 1988).

As mentioned above with respect to neopatrimonial regimes, patron–client relations, which are often the basis of politics in the third world, benefit both sides, the patron who receives the vote of the client and the client who takes advantage of the benefits provided by the patron. Furthermore, the benefits the client receives from the patron are known, certain, and immediate. Moreover, the patron, or *political entrepreneur*, as Barbara Geddes

(1994) terms it, has similar certainty that this form of exchange will enable the patron to stay in office. However, the dilemma politicians face concerns long-term reform of the economy, politics, or the broader society. As long as the reforms proposed coincide with the activities that need to be done for the politician to remain in office, reform will be delivered and will take place. But if the immediate effect of maintaining office collides with long-term reform prospects, reform will remain unrealized for years to come. Most of the time, the *politician's dilemma* inhibits the growth of state capacity and reform because politicians choose the immediate effects over long-term reform even if reform would improve their long-term credentials.

Previous scholarship treated the state as a unitary actor, behaving almost as an individual (Skocpol, 1979, 1985). Geddes (1994), however, shows that the state is composed of many different segments, such as parties, legislature, or the parliament, in each of which officials are seeking to fulfill their own rational self-interests. This self-interest includes reelection for the next term, or "maximizing their career success" (Geddes, 1994, p. 7). This dynamic heavily influences the ways in which the overall political system works. Further, administrative reform is a collective good from which everyone is likely to benefit, but organizing the reform would be irrational for individuals, and therefore the collective action problem arises as a result. The collective action problem occurs when individuals have little incentive to participate in the production of public goods; however, once the public good is provided, they will enjoy the benefit of it.

According to Robert Axelrod (1984), cooperation is possible to achieve if groups are small enough to detect *free riders*, individuals who prefer to enjoy the benefits of public goods but do not participate in the production of them. From small-scale cooperation, large-scale cooperation evolves over time because cooperation spreads to other groups as well. Geddes (1994) disagrees with this claim and states that smaller groups actually hinder the emergence of wider cooperation because of the patron–client relationship inherent in some societies. In other words, solving the collective action problem in small groups will not lead to the spread of cooperation to the larger society. In patron–client arrangements, cooperation is irrational because clients receive goods for their votes and noncooperation benefits them personally and directly.

Geddes (1994) proposes ways to deal with the politician's dilemma and potentially increase state capacity. First, civil servants should be insulated from political influence and need to be cut off from direct political involvement. Second, positions for administrative personnel have to be built on a heavily merit-based system that asks for the specialized skills the job requires. Finally, Geddes states that new agencies have to be created that are autonomous from the government, such as the Federal Reserve Bank or the Federal Aviation Administration.

The origins of *weak states* can be traced back to the global economy, as well as the colonial heritage many of these states experienced. The industrial revolution in Europe led to the development of entire regions' producing raw materials for the insatiable European market. Furthermore, the commercialization of agriculture affected more than 90% of the overall world population. Previous communal holdings were abolished, and land tenure laws were altered substantially. As a result, many people lost their land, and the old social control declined markedly. Another factor that is believed to have led to developing countries' becoming weak states is those countries' history of colonialism (Clapman, 1985; Young, 1995). Colonies were economically and culturally peripheral to the mother country. Previous local political systems were displaced by colonial administrative systems, which still form the basis of nation building today.

Despite established political institutions, former colonies remain weak states after decolonization (Migdal, 1988). Although constitutions have been developed in most of these states, those documents mostly remain papers with little enforcement power. Overall, the state remains weak and is unable to penetrate society, regulate social relations, or extract resources that would enable the state to function properly. One of the main reasons for the staying power of weak states is the persistence of patron–client relations (Powell, 1970; Scott, 1972). Peasants for the most part remain under this relationship because the patron provides insurance in case of natural disasters and poverty. The bond that ties the peasant to the patron is perhaps too strong to be destroyed.

Weak states are unable to dominate over rules of other social institutions, such as families, clans, tribes, patron–client dyads, and even multinational corporations (Migdal, 1988). Multiple sets of rules exist within society and often form gridlock with one another. Part of the reason the patron–client scenario persists is that people feel the state would be unable to provide security in case of economic hardship. People feel that this is a task a patron is better equipped and willing to deal with. On the other hand, there is also widespread resistance coming from the side of the strongmen, such as chiefs, landlords, and bosses.

Joel Migdal (1988) describes a few strategies that weak states employ for political survival, given the nature of present relations, in which the state has difficulty penetrating society. In a phenomenon termed the *big shuffle*, state leaders periodically appoint, remove, and shuffle heads of agencies and institutions that have the potential to accumulate political power, or in other words, that would have the ability to threaten the current regime. Also, in order to secure and to keep the position of the present leadership, people are appointed to key positions who have deep personal loyalties to the leader in forms of kinship, clan or tribe ties, and so forth. Further, leaders of weak states frequently resort to the imprisonment and

torture of the opposition. Each of these strategies hinders weak states from developing more complex institutions that could help them democratize. Increasing the complexity of state structure is believed to be one way to counter fast-paced development and modernization (Huntington, 1965). By artificially preventing the development of institutions, weak states close themselves off from future development.

Conclusion

Statist theories define what is understood by the state. However, this definition has been fairly ambiguous and has expanded to include political parties, pressure groups, the mass media, and so forth. As a result, frequently the term *state* has been replaced by other phrases, such as *government* or *political system*. Scholarship of statism also looks at state formation within the Western context, as well as that of developing countries. Further, statist theories also describe various functions for which the state organizes. These functions include the monopoly of warfare within set state boundaries (Tilly, 1985; Weber, cited in Gerth & Mills, 1946), taxation, the provision of social services, and so forth. Views of the state and its role in society have changed over time. Initial Western formal theories have looked at the state from the institutional standpoint. Subsequently, the society-centeredness approach, which viewed the behavior of society and interest groups as more crucial than the workings of the state, dominated. With Skocpol's works, however, the state was brought back to the fore of social inquiry. Finally, statist theories also recognize the nuances among authoritarian states and reach deeper into understanding this subset of systems. Works on the politician's dilemma (Geddes, 1994) and weak states (Migdal, 1988) have begun to provide a more balanced view of the interaction or the two-way relationship between state and society.

References and Further Readings

- Almond, G. (1988). The return to the state. *American Political Science Review*, 82(3), 853–874.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Ayoob, M. (1992). The security predicament of the third world state: Reflections on state making in a comparative perspective. In B. Job (Ed.), *The insecurity dilemma: National security of third world states* (pp. 63–80). Boulder, CO: Lynne Rienner.
- Bova, R. (1991). Political dynamics of post Communist transition. *World Politics*, 44, 113–138.
- Bratton, M., & van de Walle, N. (1994). Neopatrimonial regimes and political transitions in Africa. *World Politics*, 46, 453–489.
- Bunce, V. (1999). *Subversive institutions: The design and the destruction of socialism and the state*. Cambridge, UK: Cambridge University Press.
- Carter, G., & Hertz, J. (1972). *Government and politics in the twentieth century*. New York: Praeger.
- Clapman, C. (1985). *Third world politics: An introduction*. Madison: University of Wisconsin Press.
- Easton, D. (1981). The political system besieged by the state. *Political Theory*, 9, 303–325.
- Eckstein, H. (1960). *Pressure group politics: The case of the British Medical Association*. Stanford, CA: Stanford University Press.
- Evans, P. B., Rueschemeyer, D., & Skocpol, T. (Eds.). (1985). *Bringing the state back in*. Cambridge, UK: Cambridge University Press.
- Geddes, B. (1994). *The politician's dilemma*. Los Angeles: University of California Press.
- Gerth, H. H., & Mills, C. W. (Eds. & Trans.). (1946). *From Max Weber: Essays in sociology* (see especially pp. 77–128). New York: Oxford University Press.
- Hillier, T. (1998). *Sourcebook on public international law*. London: Routledge. Retrieved May 15, 2009, from <http://books.google.com/books?id=Kr0sOuIx8q8C> (see especially pp. 201–202).
- Huntington, S. (1965). Political development and political decay. *World Politics*, 17(3), 386–430.
- Krasner, S. (1984). Approaches to the state: Alternative conceptions and historical dynamics. *Comparative Politics*, 16, 223–246.
- Lane, F. (1958). Economic consequences of organized violence. *Journal of Economic History*, 18(4), 401–417.
- Lauterpacht, H. (1958). *The development of international law by the International Court*. New York: Praeger.
- Linz, J. J. (2000). *Totalitarian and authoritarian regimes*. Boulder, CO: Lynne Rienner.
- Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and post Communist Europe*. Baltimore: Johns Hopkins University Press.
- Malloy, J. M. (1977). Authoritarianism and corporatism in Latin America: The modal pattern. In J. M. Malloy (Ed.), *Authoritarianism and corporatism in Latin America* (pp. 3–19). Pittsburgh: University of Pittsburgh Press.
- Mason, T. D. (2004). *Caught in the crossfire*. Lanham, MD: Rowman & Littlefield.
- Migdal, J. (1988). *Strong states and weak societies*. Princeton, NJ: Princeton University Press.
- Miliband, R. (1969). *The state in capitalist society: The analysis of the Western system of power*. London: Quarter Books.
- Montesquieu, C. (1989). *The spirit of the laws* (A. Cohler, B. C. Miller, & H. S. Stone, Eds. & Trans.). Cambridge, UK: Cambridge University Press.
- Nordlinger, E. A. (1981). *On the autonomy of the democratic state*. Cambridge, MA: Harvard University Press.
- Nordlinger, E. A., Lowi, T. J., & Fabbrini, S. (1988). The return to the state: Critiques. *American Political Science Review*, 82(3), 875–901.
- O'Donnell, G. (1978). Reflections on the patterns of change in the bureaucratic authoritarian state. *Latin American Research Review*, 13, 3–8.
- Powell, J. D. (1970). Peasant society and clientelist politics. *American Political Science Review*, 64, 411–425.

- Putnam, R. (1973). The political attitudes of senior civil servants in Western Europe. *British Journal of Political Science*, 3, 257-297.
- Ranney, A. (1966). *The governing of men*. New York: Holt, Rinehart & Winston.
- Reno, W. (1997). War, markets, and the reconfiguration of West Africa's weak states. *Comparative Politics*, 29(4), 493-510.
- Scott, J. C. (1972). Patron-client politics and political change in Southeast Asia. *American Political Science Review*, 66, 68-90.
- Skocpol, T. (1979). *States and social revolutions: A comparative analysis of France, Russia, and China*. Cambridge, UK: Cambridge University Press.
- Skocpol, T. (1985). Bringing the state back in: Strategies of analysis in current research. In P. B. Evans, D. Rueschemeyer, & T. Skocpol (Eds.), *Bringing the state back in* (pp. 3-43). Cambridge, UK: Cambridge University Press.
- Tilly, C. (1985). War making and state making as organized crime. In P. B. Evans, D. Rueschemeyer, & T. Skocpol (Eds.), *Bringing the state back in* (pp. 169-191). Cambridge, UK: Cambridge University Press.
- Tilly, C. (1990). *Coercion, capital, and European states*. Malden, MA: Blackwell.
- Truman, D. (1951). *The governmental process*. New York: Knopf.
- Watkins, F. (1968). *International encyclopedia of the social sciences*. New York: Macmillan.
- Young, C. (1995). The heritage of colonialism. In J. W. Harbeson & D. Rothchild (Eds.), *Africa in world politics: Post-cold war challenges* (pp. 23-30). Boulder, CO: Westview Press.

12

DEPENDENCY AND DEVELOPMENT

FRANK A. BOYD JR.

Illinois Wesleyan University

What explains the divergent development trajectories of different countries and societies? Why has the developing world failed to close the gap with the industrialized countries? This chapter reviews the literature by scholars in the Marxist tradition that addresses these and other questions of development, beginning with the writings of Karl Marx, continuing through the work of Vladimir Lenin, and ending with the more recent work of the dependency theorists. This literature is vast. And like the writings of Marx himself, there are often competing or even opposing theoretical arguments from scholars considered part of this tradition. Taken together, this literature forms an enduring theoretical counterpoint to the liberal economic theory that has informed generations of policymakers from both developed and developing countries.

Marxist, neo-Marxist, and dependency explanations of development differ from traditional liberal economics in the importance of noneconomic factors for explaining development. Liberal economists emphasize that countries can best promote development by limiting social and political concerns from policy making and by integrating their national economies into the global market. In response, the literature featured in this chapter suggests that development can be explained only through an understanding of the social and political underpinnings of policy making and that self-interested countries should view economic interdependence as inherently asymmetrical. Neo-Marxist

and dependency theorists argue that capitalism distributes the benefits or profits from economic production in an uneven fashion, so specific groups or classes of individuals benefit more than others in a domestic economy. This logic is extended to the international system, where the structure of global markets ensures that wealth becomes concentrated in a small group of countries at the expense of a larger group of developing countries.

This chapter now turns to Marx and the theoretical beginnings of this approach, followed by a brief overview of Lenin's theoretical revision and extension that served as a benchmark for the neo-Marxists of the early 20th century. The last half of the chapter attempts to impose some order on the sprawling literature referred to as *dependency theory*. These scholars borrowed, often selectively, from orthodox Marxism to explain the development fortunes of the developing world and the relationship of these countries to their industrialized counterparts in the North.

Marx and Lenin

Karl Marx (1777) and the scholars who followed him characterize capitalist economies as inherently conflictual and exploitative. In his critique, Marx described how the dynamics of capitalism provide a powerful force for producing and accumulating wealth. Adam Smith (1776/1994), David Ricardo (1817), and the liberal economists that followed

them explained that capitalist markets are self-regulating. But over the long term, Marxists argue, capitalist markets do not have a natural, self-regulating equilibrium. Instead, capitalist systems are subject to cycles that will impose great costs on societies and will increase in severity as capitalism matures. Workers, rather than producers, will bear the economic and social costs of these market fluctuations. Marx argued that workers ultimately would reject capitalism in favor of socialism by way of a popular revolution. Of course, a worldwide socialist revolution has not materialized, but scholars working in the Marxist tradition have preserved the central assumptions of Marx to explain political and economic development in the least developed countries (LDCs), and some continue to call for a Marxist revolution.

Marx's writing was wide-ranging, and his scholarship has spurred significant, and often divergent, scholarship. His explanation of capitalist development focused on the productive process in which economic activity produced surplus value, or profits, that accrued to the bourgeoisie, or owners of capital, instead of to workers. Marx explained that the creation of surplus value is necessary for capital accumulation, which is an important determinant of mid- to long-term economic development. Because workers, by definition, are excluded from ownership, they also are excluded from reaping the true benefits of their labor.

Marx employed a specific analytical approach, which he described as a materialist conception of history, or *historical materialism*. From this perspective, societies progress over time through different modes of production determined by the tools, natural resources, technology, and other productive resources available to them. Marxist scholars disagree on the number of specific modes, but the modes begin with tribal societies, continue through feudalism, and end at the highest level of development, the capitalist mode of production. At each mode, or stage, societies organize themselves differently and develop specific and economically contingent ways of dividing the surplus of their productive activity. This process yields a specific social and political superstructure that rests on the mode of production and is specific to a society's stage of economic development. Feudal societies, for instance, had social and political institutions dramatically different from those of the societies in the newly industrializing countries of western Europe in the late 19th century. Marx's analysis of early industrial Europe focused on the transition from a feudal political economy to a capitalist system. The most important implication of Marx's insight is that political relations are contingent on and will respond to changes in productive activity. It is on this nexus—the recursive relationship between economic and political development—that neo-Marxist and dependency theorists focus their scholarship.

Robert Gilpin (1987) notes that Marx's critique also identified three internal contradictions of capitalist economic systems that would cause tensions between the economic and political development trajectories of countries.

The first is the *law of disproportionality*. Liberal economics assumes that markets tend naturally toward equilibrium, with the supply of goods produced either rising or falling to meet demand. The prices of goods will also fluctuate in response to market signals, ensuring that capitalist economic systems will rarely suffer from severe disequilibrium. Instead of equilibrium, Marx proposed the law of disproportionality, which stated that the productive efficiency of capitalism will instead trigger increasingly severe economic fluctuations, leaving markets with an oversupply of goods or workers (consumers) who lack the resources to buy them. When the social and political fallout from disproportionality combines with the other negative consequences of capitalism, social upheaval and political revolution will result.

The second law emphasized by Marx is the *law of capital accumulation*. Capitalist markets provide opportunities for producers to efficiently reap profits and to accumulate the surplus value, or wealth, of economic activity. Over time, the internal dynamics of capitalism tend to concentrate wealth in the hands of a few, which in turn can affect the productive and consumptive capacity of domestic economic systems. This concentration of wealth becomes problematic if the capital is not invested in ways that generate additional economic activity. For example, producers may choose not to invest their capital but to transfer the profits overseas or use their capital for consumption. Thus, capitalist systems can be denied investment that is required for continued economic growth. Marx argued that the concentration of capital would ultimately breed social and political unrest among workers that could destabilize countries or even trigger a popular revolution.

The third and final law of capitalism emphasized by Marx is the *long term decrease in profits* realized by producers. The discipline of competitive markets encourages producers to protect their profit margins by increasing the efficiency of their operations, by reducing labor costs, by incorporating technology, or by other means of depressing wages. Workers, faced with higher rates of unemployment or reduced purchasing power, drive down the price of goods and further depress the profits of producers. Again, Marx predicts that political and social costs of this dynamic unrest will eventually trigger popular protest to confront the costs of capitalism.

Marx's early work provided a sophisticated critique of capitalism, but he found that his theoretical framework did not fit very well when he turned his attention to the countries of Asia, where the feudal and precapitalist characteristics that he observed in Europe did not exist. Moreover, what is called the Asiatic mode of production did not have the class conflicts that Marx observed in western Europe, which left Asian countries without the dialectical class conflict that can propel a country's development trajectory.

Nearly 40 years passed between Marx's work and the first publishing of Lenin's (1939) major work, *Imperialism*,

which undertook the task of refining Marxism to explain the rapid diffusion, expansion, and persistence of capitalism. There was much to explain. Capitalism had expanded and flourished between 1870 and 1917 and had suffered none of the systemic problems proposed by Marx. *Imperialism* provided a description of how capitalist countries avoided the political upheaval predicted by Marx through imperial expansion of their economic and political power. In short, Lenin extended Marx's critique of capitalism from domestic political economy to the global sphere.

In the late 19th and early 20th centuries, the European economic powers acquired vast colonial holdings in Asia, Africa, and Latin America. These colonies helped extend the reach of capitalism by providing sources of raw materials, markets to absorb the overproduction of manufactured goods, and capital investment opportunities for European firms. The financial capital that was made available by industrialization allowed the European powers to exploit the workers and extract the natural resources of their far-flung colonies. In a theme that would be the focus of the dependency theorists some years later, Lenin explained that the imperial powers used their economic might to co-opt the political and economic elites native to the colonies.

Moreover, the development of an imperial system allowed industrialized countries to secure new markets for their manufactured goods. As firms became more efficient, the supply of goods soon exceeded the domestic demand for their goods. Lenin explained how imperialist expansion allowed the industrialized countries to export their manufactured goods to the captive markets of their colonies, thereby muting some of the negative aspects of capitalism that had been emphasized by Marx.

For Lenin, the imperialism that he observed in the early 20th century represented a more advanced form of capitalism but was nevertheless a system that in the long term would suffer from systemic instability. Imperial capitalism generated rapid but uneven development, a theme cited by Marx in his domestic analysis of capitalism. Once the imperial powers had expanded capitalism to include the entire developing world, Lenin predicted that the negative consequences of capitalism would trigger conflict among the imperial powers. Moreover, the economies of the developing world—propelled by the forces of capitalism—would eventually expand to compete with the imperial powers. These economic conflicts among and between capitalist countries would eventually result in political competition, armed conflict, and the eventual demise of the capitalist system.

Dependency Theory

Dependency theorists returned to the themes of Marx in the mid- to late 20th century, and they presented a collection of approaches to the political economy of developing countries. Mainstream economic thought of the

time clustered around *modernization theory*, whose foremost proponent, W. W. Rostow (1960), suggested an explanation of economic development that borrowed the Marxian notion of stages of growth. These scholars saw the development of all countries proceeding along a similar trajectory and believed that differing levels of development simply reflected the position of countries along this set developmental path. This perspective argued that the economic and political factors that had been important for explaining the rapid development of the countries in the North—resource endowments, labor inputs, technology, and investment capital—could similarly drive the development of LDCs. In fact, the most rapid path to economic development was rapid integration into the global political economy.

But dependency theorists argued that the characteristics of LDCs and their position in the global political economy strongly conditioned their prospects for growth. In fact, the relative poverty of the developing world could not be explained as a function of their relative isolation from the global political economy, but instead could be explained by the manner in which they were integrated into the global capitalist system. To explain the shortcomings of Rostow's approach, dependency theorists drew on the whole panoply of Marxist and neo-Marxist thought, while adding some theoretical refinements that better explained the pattern of development they observed in the 1960s and 1970s. For instance, early Marxists concentrated primarily on the relations of production, whereas the dependency theorists placed more emphasis on the structure of unequal exchange.

The dependency literature can be divided into two essential approaches. The first group posits the “development of underdevelopment” and was advanced by Paul Baran in the 1950s and Andre Gunder Frank in the mid-1960s. These early dependency scholars focused principally on the international dimension of dependent development. Baran, Frank, and others certainly incorporated domestic political, social, and economic factors into their analysis, but they devoted little analytical space to explaining domestic actors and institutions. This group of scholars also agrees on the need for a worldwide Marxist revolution to overthrow the capitalist system.

A second theoretical cluster eschews the revolutionary political dimension of the early dependency theorists and Immanuel Wallerstein. The scholars of this group reject the necessity of a socialist revolution and instead suggest that LDCs can harness the economic power of capitalism to promote development in the periphery (defined below). This literature also devotes analytical space to domestic politics and elaborates a finely grained explanation of the relationship between domestic political factors and the international political economy. Overall, these scholars attribute more agency to domestic political actors on the periphery and argue that in some cases, the processes of dependency and development can coexist in LDCs.

The “Development of Underdevelopment”

Scholars often cite Baran (1967) as the first of the neo-Marxist scholars of the dependency school. Baran's work incorporated many of Marx's basic assumptions about political economy and refined them to better explain the development of LDCs in the mid-20th century. He began with the basic observation that the capitalist system rests on an exploitative relationship between the industrialized countries of the “center” and the developing countries of the “periphery.” For Baran, the functioning of a capitalist economy requires exchange relations that produce more rapid economic growth in one group of countries at the expense of development in the second group of countries. Thus, the functional relationship between development and underdevelopment is a requisite characteristic of capitalism, and the prospect of rapid economic growth in the LDCs actually threatens the profitable functioning of capitalism and the economic fortunes of industrialized countries.

This center-periphery structure of the international political economy takes Marx's class-based theory of development and extends it to the international system. An international division of labor segregates the countries of the periphery into a functional role in which they produce primary commodities to be used in the manufacturing industries of the center. In the early 20th century, Lenin emphasized the imperialistic exploitation of colonies by the colonial powers. Baran's (1967) work sharpened this theoretical insight and observed that the political economy of LDCs had changed very little since Lenin's time, with agriculture and mining remaining the dominant economic activity.

Baran's (1967) most significant theoretical contribution is a revision of Marx's theory of surplus value. Baran suggests that the problems of underdevelopment can all be traced to the underutilization of economic surplus. Baran proposed a three-part typology of surplus: actual surplus, potential surplus, and planned surplus. *Actual surplus* refers to the difference between production and consumption; *potential surplus* indicates what societies could produce and what they need to consume; *planned surplus* indicates a society's planned production and optimal consumption. Baran conceded that measuring both potential and planned surplus would be difficult in practice, especially given the difficulty in establishing an indicator of a country's actual surplus. Still, these theoretical concepts emphasized the structural position of LDCs in the world political economy and how the dominance of monopoly firms in developing countries weakened the bargaining positions of LDC governments. Taken together, these structures make it difficult, if not impossible, for LDCs to accumulate the capital that is necessary for long-term development.

If the profits that are earned in LDCs are not reinvested there, where do they go? Rather than investing their profits in higher-value-added manufacturing, Baran

explains, the elites in LDCs tend to replicate the consumption patterns of their counterparts in the industrialized center. The purchase of imported luxury goods and consumer durables contributes little, if anything, to LDC economies. In addition, these same elites often take surplus capital and deposit it in overseas banks for protection against market fluctuations or political unrest in their home country. Hence, developing countries are never able to provide the capital accumulation that is necessary for economic development.

Frank (1967) extended many of Baran's observations about the underdevelopment of LDCs, beginning with the assertion that the industrialized center realizes more relative gains in its economic relationships with the developing countries of the periphery. Instead, in Frank's view, development and underdevelopment are two interrelated parts of the international capitalist system, which is divided into the *metropole* of developed countries in the center and the LDCs of the periphery. The capitalist system binds together the economic and political fortunes of the core and periphery in an exploitative system of exchange. The very design of the capitalist system produces economic surplus that is drawn out of the periphery to enrich the metropole countries. To describe this dynamic, Frank coined the pithy phrase “the development of underdevelopment.” He even argued that this exploitative system replicated itself within developing countries, where industrialization can be found in the urban centers at the expense of the rural poor, who remain subjugated to the political and economic dominance of the urban elite.

Frank's (1967) work differs from classical and neo-Marxists in one significant respect: He rejected Marx's historical modes of production. Instead, Frank contends that the economic systems of LDCs have long been capitalist in structure and that countries do not pass through a sequence of production modes that ultimately ends in capitalism. Empirically his analysis concentrates on Latin America, whose countries have been integrated into the global capitalist system since their political independence in the early to mid-19th century. Frank observes that there exist few observable differences in the economic structures of Latin American countries between the time of Hernando Cortés and the present. Frank argues that these countries did not struggle to leave behind a feudal economy, polity, and society but instead were inserted into the capitalist economy. From this perspective, countries are not underdeveloped because they have yet to pass through the successive stages of economic development to reach capitalism; rather their underdevelopment is a necessary by-product of the development of the center.

Modern World-Systems Theory

A final theoretical refinement and extension of Baran's and Frank's work can be found in modern world-systems (MWS)

theory, of which Wallerstein (1974) is the foremost proponent. MWS scholars built on the work of early *dependistas* (dependency theorists), and they also rejected some of the core assumptions of Marx. For instance, where Baran, Frank, and other early dependency theorists paid scant attention to the domestic political economy, the MWS theorists formally asserted that the analytical focus should be on the international system, whereas Marxism concentrates on the domestic class struggles over production and the surplus value that comes from it, and Wallerstein posits an international system of states divided into center, semi-periphery, and periphery.

This class-based system places an economically dominant group of countries at the core and a much larger group of underdeveloped countries on the periphery, all of which together constitute a functional whole. Of course, Marxist scholars since Lenin have posited a center-periphery capitalist structure, but Wallerstein added a third category of states, the *semi periphery*, to the traditional dualism of Marxist and neo-Marxist theorists. The political systems of the semi-periphery allowed them to marginally resist the political power of the core and to capture some of the economic surplus that would otherwise be transferred out of their country. The addition of the semi-periphery also made Wallerstein's world capitalist system more closely reflect the economic and political development of countries in the mid-20th century, when a number of countries, such as Brazil, Korea, and Taiwan, had significant manufacturing sectors that competed with industries in the center.

MWS theorists argue that the same dynamic capitalist forces that drive economic development in the core countries also produce underdevelopment in the countries of the periphery. In an extension of Frank's "development of underdevelopment," this argument holds that capitalism will not produce development in the countries of the periphery over the long term. Instead, economic development of the core necessitates a concomitant underdevelopment at the periphery. Rather than being loosely joined, the countries of the core and the periphery are mechanistically linked so that the economic surplus of the dependent periphery is transferred to the core. Wallerstein makes this argument in its most pure form by stating that a country can have no national development that is independent of the larger system within which it resides. Wallerstein and Frank diverge from classical Marxist thought by refuting Marx's historical modes of production. They counter with the claim that the countries of the periphery will never experience the capitalist development of the center but are fixed in a permanent state of underdevelopment.

Dependency Theory Without Revolutionary Marxism

The origins of *structuralism* actually predate the early dependency theorists, whose roots are in the 1950s, but

structuralism came to the fore in 1964 with the publication of the report titled *Towards a New Trade Policy for Development*. This report and the work of the U.N. Economic Commission for Latin America became identified with structuralism and the work of Raúl Prebisch (1950). Structuralists differ from the early dependency theorists in their rejection of the Marxist and neo-Marxist call for a revolutionary overthrow of capitalism. In fact, structuralism has little of the Marxist ideological approach found in much of the other dependency literature.

Nevertheless, structuralism shares many theoretical assumptions with the other dependency theorists, beginning with an understanding that global capitalism establishes unequal terms of trade between the center and periphery. Many of the same factors that expanded international trade in the 19th century allowed late developers such as Germany and Japan to combine to perpetuate or even widen the developmental gap between industrialized and developing countries. For example, capitalism requires ongoing technological innovation to reach higher levels of productivity and to generate ever increasing profits. Structuralists join other dependency theorists in arguing that technological innovation does not always spur economic growth and development, especially for the commodity-producing countries of the periphery.

Because they are "late, late developers," LDCs remain the principal producers of primary commodities, whereas high-value-added industrial production is found in the center.

In industrialized countries, the competitive forces of capitalism compel producers to introduce technological innovations to increase the value of manufactured goods by lowering production costs and increasing their efficiency. Workers become displaced in this process, and in industrialized countries, investment flows are sufficient to provide other employment opportunities for these workers. Thus, the competitive destruction of capitalism provides for higher levels of efficiency while still providing higher-wage employment for workers.

The structural characteristics of LDCs, however, mean that the benefits of technological investment are less apparent and potentially costly for several reasons. First, technological innovation is almost always imported from the core and directed at agricultural production or very light, low-value-added manufacturing. Second, most dependency theorists explain that the amount of capital investment available for developing countries is much lower than in the industrialized core. This stems, in part, from the generally low savings rates in poorer countries. As Baran argued, the profits that might be used for domestic investment are either repatriated to the core or consumed by the local elites, whose consumption patterns more closely mirror those of the industrialized core than those of their fellow citizens in the periphery. To be sure, technological innovation and capital investment in agriculture can increase efficiency, but there is a limit to the innovation

possible in the production of primary commodities. Even if they still have jobs, workers are left with depressed wages, and displaced workers face a labor market with few alternative employment opportunities.

Beyond the considerable overlap with other dependency theorists, structuralists offer an important theoretical contribution regarding the declining terms of trade faced by countries on the periphery. Structuralists argue that there is a long-term decline in the terms of trade for developing countries, which depend on primary exports, in relation to the industrialized core, whose manufacturing sector generates products with ever increasing value. Over time, the profits for primary commodities do not increase at the same rate as do the profits for higher-value-added goods. This inequity between export revenue and the price of exports creates imbalances in a country's balance of payments.

The evidence for a long-term decline in the terms of trade is mixed, with many developing countries reporting periods of decline in their terms of trade, interspersed with periods of rapid improvement. This pattern points to another important contention of structuralists regarding the fluctuations of commodity markets. The instability of commodity markets makes for much sharper and unpredictable business cycles in developing countries. This is especially true for countries that depend on a small basket of primary exports, because a sharp drop in the profits for one commodity can have very serious economic implications. Moreover, the demand for primary exports is externally determined and contingent on the appetite for those commodities in the industrialized North. Because they are unable to predict or anticipate these market fluctuations, countries are unable to craft long-term planning for development. Again, the structure of the international market has a differential impact on the core and the periphery.

Of all the scholarly work reviewed in this chapter, structuralism has had the most direct impact on policy making in LDCs. In response to plummeting commodity prices and unavailability of manufactured goods, a number of LDCs have attempted *import substitution industrialization*. Countries on the periphery erected tariff and nontariff barriers to protect infant industries that focused on producing goods for domestic consumption. In the short term, some countries rapidly developed manufacturing sectors, but many of these sectors collapsed in the longer term. The domestic markets of LDCs, even in larger countries such as Brazil, proved to be too small to support a profitable manufacturing sector. Also, the same protection that kept multinational corporations out of the markets of larger LDCs also removed any incentives for domestic firms to improve quality or efficiency of production. For this reason, Brazil, Argentina, and Mexico eventually supported bloated, inefficient industries that faced no significant competition from international firms. Few countries pursued import-substitution industrialization as a development strategy after the 1980s.

Dependent Development

Fernando Cardoso and Enzo Faletto first published their work, *Dependency and Development in Latin America*, in 1967, but an English translation did not appear until 1979. Although they are grouped with the dependency theorists, many of their theoretical assumptions are at odds with the other scholars in the school. Cardoso and Faletto (1979) observed that the predictions that flowed from Baran, Frank, and other early dependency theorists did not correspond with the development experiences of many developing countries. A number of countries on the periphery had large and growing industrial sectors. By the late 1960s, the countries of the Southern Cone in Latin America had burgeoning industrial sectors, and several countries of East and Southeast Asia boasted an expanded manufacturing sector.

The economic growth and development in these countries—then referred to as the *newly industrializing countries*—called into question the determinism of the early dependency theorists. Why, they asked, did industrialization proceed in some developing countries and not in others? Cardoso and Faletto place great emphasis on the domestic political dimension of dependency. Specifically, they argue that LDCs can effectively intervene on behalf of national capital and that international capital can be encouraged to invest in domestic economies for manufacturing and for nonexport production. Nevertheless, Cardoso and Faletto note that the bargains struck by elites in developing countries often do not result in a broad distribution of wealth among the broader population of developing countries. The maldistribution of wealth creates both economic and political challenges for LDC governments and also creates a tendency to adopt nondemocratic or even what Guillermo O'Donnell (O'Donnell, Schmitter, & Whitehead, 1986) calls *soft authoritarian* forms of governance.

Peter Evans (1979) followed Cardoso and Faletto in defining a more autonomous role for the state in developing countries, and in explaining in more detail the domestic political dimension of development. According to the early dependency theorists, the actors and institutions of peripheral states had very little agency in determining the prospects for economic development and simply responded to the structural characteristics of global capitalism. The early dependency theorists either ignored or reified the preferences of these different actors, and they did little to explain how or why bargains emerged among them. Evans argued instead that states in some instances play a central role in organizing the alliances between domestic political and economic actors and the international capitalists who would invest in their countries. Evans also provided much more insight into the relationship between the state, international capital, and the local elite and explained how this “triple alliance” could promote economic development and capital accumulation in developing countries.

Evans's empirical analysis centered on the development experiences of Brazil, where the industrial sector grew considerably during the first half of the 20th century, but only with the consistent pressure of the local elite and the state. His study explains that the state played a key role in securing an agreement between local and international capital but also found it very difficult to strike a bargain that would provide for a broader distribution of the profits. Evans also explains that the experience of Brazil could very well be extended to the other countries of Wallerstein's semi-periphery.

Conclusion

The critique of capitalism set forth by Marx explains how the powerful forces unleashed by global capitalism would transform societies from feudalism to capitalism. He observed how the forces of capitalism spread from Great Britain to Europe and believed that ultimately capitalism would revolutionize the productive modes of the entire world. The internal contradictions of capitalism, however, would create a conflict between workers and the owners of the means of production. The conflict between these two classes, which Marx described as a dialectical struggle, would eventually yield to socialism. Most important, however, is the claim that capitalism would eventually diffuse to every country in the global economic system.

Lenin built on the theoretical insights of Marx, viewing the expansion of capitalist markets and production as essential for moving countries from their historic modes of production, such as feudalism, for instance, to worldwide socialism and ultimately Communism. In many ways, Lenin's argument stood opposite many dependency theorists. Lenin believed that capitalism would promote the development of the periphery rather than retard its chances for long-term economic growth. Conflict would arise when the imperial countries were challenged by their newly developed colonies. Thus, Marx and Lenin theorized that global capitalism promoted the development of LDCs, although that development might proceed in an uneven fashion.

Dependency theorists, although departing from Marxist orthodoxy in some respects, draw on the Marxist tradition to explain economic and political development. Nearly every summary of dependency theory emphasizes the heterogeneity of this work, which makes the grouping of its literature somewhat difficult. This chapter posits two groups. The first group, referred to as the *early dependency theorists*, emphasizes the need for or inevitability of a socialist revolution, and the second group concentrates on the possibility of dependent development.

Even though the dependency literature covers significant and at times contradictory theoretical space, nearly

all dependency theorists share a few common assumptions: The prospects for development in LDCs are externally determined or strongly conditioned by factors external to the developing countries; the countries of the center benefit from unequal exchange with the periphery; and even in the absence of national development of LDCs, the local elite often benefits from the investments from industrialized countries. Baran suggests that the capitalist development of countries in the industrialized center formed a functional economic link with the LDCs of the periphery. Frank extends this argument by arguing that global capitalism fueled the development of underdevelopment of the periphery. This concept, it is worth noting, contradicts the predictions of Marx and especially Lenin. Wallerstein provides an extreme example of how these early dependency theorists paid scant attention to the domestic political economy of development in LDCs.

As noted earlier in this chapter, the later dependency theorists perceive a less deterministic relationship between developed and less developed countries. Cardoso and Faletto also see the development prospects of LDCs as strongly conditioned by global capitalism, and they attempt to explain the uneven pattern of development among developing countries. Evans extends this perspective in his explanation of Brazil's industrialization by suggesting a triple alliance of multinational corporations, the state, and local capital (or bourgeoisie). Rather than presenting a deterministic model, Evans suggests that development policies emerge from bargaining, and in the case of Brazil, a path of dependent development emerged. In contrast to the earlier literature, these scholars devote more space to analyzing the structure and function of domestic actors in LDCs.

There remains a significant gap between the level and trajectory of development among LDCs and that of the industrialized center. But the notion of the developing world as a single, relatively homogeneous group of countries is no longer accurate or analytically useful. As a result, scholars in the 21st century are less likely to employ the analytical approach set forth by the dependency theorists. It is interesting that scholars and especially journalists are more likely to invoke some of Marx's perspectives to explain the increased frequency of global financial crises. To be sure, the conditioning effects of global capitalist markets will continue to be of interest to scholars interested in explaining the variations in development among groups or classes of countries.

References and Further Readings

- Amin, S. (1977). *Imperialism and unequal development*. Sussex, UK: Harvester Press.
- Baran, P. (1967). *The political economy of growth*. New York: Monthly Review Press.

- Brewer, A. (1980). *Marxist theories of imperialism: A survey*. Boston: Routledge & Kegan Paul.
- Caporaso, J. A. (Ed.). (1978). Dependence and dependency in the global system [Special issue]. *International Organization*, 32(1).
- Cardoso, F. H., & Faletto, E. (1979). *Dependency and development in Latin America*. Berkeley: University of California Press.
- Chilcote, R. H. (1982). *Dependency and Marxism: Toward a resolution of the debate*. Boulder, CO: Westview Press.
- Cox, R. W. (1979). Ideologies and the new international economic order: Reflections on some recent literature. *International Organization*, 37, 157-188.
- Dos Santos, T. (1970). The structure of dependence. *American Economic Review*, 60, 231-236.
- Emmanuel, A. (1972). *Unequal exchange*. New York: Monthly Review Press.
- Evans, P. (1979). *Dependent development: The alliance of multi-national, state, and local capital in Brazil*. Princeton, NJ: Princeton University Press.
- Frank, A. G. (1967). *Capitalism and underdevelopment in Latin America*. New York: Monthly Review Press.
- Gilpin, R. (1987). *The political economy of international relations*. Princeton, NJ: Princeton University Press.
- Heilbronner, R. L. (1980). *Marxism: For and against*. New York: W. W. Norton.
- Lenin, V. (1939). *Imperialism: The highest stage of capitalism*. New York: International Publishers.
- Marx, K. (1977). *Karl Marx: Selected writings* (D. McLellan, Ed.). Oxford, UK: Oxford University Press.
- O'Donnell, G., Schmitter, P., & Whitehead, L. (1986). *Transitions from authoritarian rule: Comparative perspectives*. Baltimore: Johns Hopkins University Press.
- Prebisch, R. (1950). *Economic development in Latin America and its principal problems*. New York: U.N. Department of Economic Affairs.
- Ricardo, D. (1817). *On the principles of political economy and taxation*. London: J. Murray.
- Rostow, W. W. (1960). *Stages of economic growth: A non communist manifesto*. Cambridge, UK: Cambridge University Press.
- Smith, A. (1994). *An inquiry into the nature and causes of the wealth of nations*. New York: Modern Library. (Original work published 1776)
- Wallerstein, I. (1974). *The modern world system*. New York: Academic Press.
- Warren, B. (1973). Imperialist and capitalist industrialization. *New Left Review*, 81, 3-44.

13

CIVIL WARS

M. ISRAEL STEPHENS

University of North Texas

Since the end of World War II, there have been 5 times as many civil wars as interstate wars and at least 5 times as many deaths due to civil wars as due to interstate wars (Singer & Small, 1994). For various reasons, the relative peace among the members of the state system did not seem to elicit a similar effect within the members of the state system during the six decades after World War II. It is perhaps not surprising then that scholarly literature on civil wars has grown substantially—and substantively—in the 21st century. Scholars ostensibly accept that civil wars are social phenomena distinct from interstate wars, which implies that civil wars likely have causes, correlates, and outcomes that are substantively different from the causes, correlates, and outcomes of interstate wars.

This chapter discusses the major theoretical contributions and controversies related to the civil war research program. First, civil wars are defined conceptually, then operationally. Second, major theories focusing on the causes of civil wars are discussed, along with future directions and policy implications. The chapter ends with a summary and a brief note on the current state of civil war research. While the discussions herein are by no means exhaustive, they tend to focus on the most notable civil war theories at this time.

What Is Civil War?

Civil war is defined in two ways: conceptually and operationally. The latter definition is subordinate to the former in that the latter attempts to make the former definition empirically useful. As will become clear, the disagreements among scholars about operational definitions can be traced to divergent opinions about what civil wars are conceptually. It is thus necessary to commence with the conceptualization of civil war prior to discussing alternative operational definitions.

Civil Wars Versus Interstate Wars

The first step is to elucidate what makes a war a *civil* war. Simply, civil wars are those fought between or among disputants within a single country, in contrast to wars fought between or among disputants of different countries. But to characterize civil wars as different from interstate wars in this manner is to miss the truly consequential differences between the two types of wars. A twofold trend of civil wars not found in interstate wars makes civil war a conceptually different type of war.

First, civil wars are fought between disputants of relatively unequal political stature. States ostensibly have the

sole legitimate authority to use coercive power within the country. All other subgroups are therefore politically inferior to the state, which renders the state inimical to subgroups' autonomy and the subgroups' autonomy inimical to the primacy of the state. That is, when subgroups desire more autonomy—such as when the Kurds attempt to dissociate themselves from the Turkish state—the subgroups begin to pose an existential threat to the state's primacy in that territory. Likewise, increased state vigilance manifestly threatens the ability of subgroups to govern themselves. For instance, China's actions in Tibet in 2008 transferred much of Tibet's local authority to the state apparatus in Beijing. In both cases, the state is politically superior to the outgroup. This contrasts with interstate wars, which involve disputants who are—theoretically—political equals. That is, no one state is legally subordinate to another. Canada has no legal right to make Lesotho's laws, nor does Lesotho have any right to impose laws on China. Were these states ever to war with each other, there would be no rebel group.

This distinction in the relative political (in)equality of the disputants in civil and interstate wars arises because of the structural contexts within which the two types of wars occur. As mentioned, an established civil hierarchy exists whereby states can force compliance legitimately. It is this hierarchy that rebels seek to deconstruct or reconstruct and which the states wish to preserve or augment. Conversely, the international system is void of a formal, constitutional order (Waltz, 1959). It is therefore argued that interstate wars likely occur as a result of disagreements or ambiguities about what states' obligations are to one another, given that they are political equals and are not inherently subordinated to one another (Waltz, 1959).¹ The second critical distinction between civil and interstate wars, then, is that civil wars occur in a context of order whereas interstate wars occur in a context of anarchy.

Although counterintuitive, that civil wars occur where there is order does not correlate with less chaos. Indeed, some posit that civil wars tend toward more chaos than do interstate wars precisely because order preexisted (e.g., Vazquez, 1993). In this view, chaos erupts because disruption of the extant order leads to a disruption of the sociopolitical paradigm by which all citizens and groups of citizens interact. On the other hand, there is no sociopolitical paradigm to be corrupted on the international stage; the only type of war affecting the macro sociopolitical paradigm of the international system would be world wars (Vazquez, 1993).

Empirical data tend to validate the idea that civil wars have a relatively more violent disposition. Mason (2004) estimates that civil wars produced 170,000 battle deaths per year in the first five decades after World War II. These deaths, according to Singer and Small's (1994) *Correlates of War Project*, accounted for 64 out of every 100 battle deaths suffered (Mason, 2004). The 1997 update of *Correlates of War* shows the same trend, with more battle deaths attributed to civil wars than to interstate wars

(Sarkees, 2000). Even these data obscure the differences in brutality. For example, scholars observe that civil war casualties tend to include a relatively high percentage of noncombatants (Holsti, 1996) as states and rebels target civilians for a variety of reasons (Kalyvas, 2006; Mason, 2004; Mason & Krane, 1989). Moreover, if it is true that civilians become targeted only after they are convincingly labeled "barbarians" (Salter, 2002), then wars in which political unequals engage in combat ostensibly would be the wars most likely to produce civilian casualties. That is, the task of labeling the enemy a barbarian within the civil war context appears much simpler than the task a state would have in labeling another state barbarous. And if such labeling is a virtually necessary condition of civilian casualties, then it follows that civilian casualties are more likely to occur in civil wars than in interstate wars. The point is this: Not only do civil wars substantively differ from interstate wars, but logic and evidence suggest that these differences explain why civil wars tend to take on a decidedly more violent disposition, including the tendency to produce substantial civilian casualties.

Civil Wars Versus Other Types of Internal Political Violence

On distinguishing civil wars from interstate wars, the task remains to discriminate civil wars from other forms of domestic mass behavior and political violence. One characteristic is that civil wars involve collective action (Lichbach, 1998; Mason, 2004). For collective action to occur, individuals face a paradox they must resolve in order to commit to group goals (Lichbach, 1998; Olson, 1971), in this case rebellion. Briefly, the benefits of a successful rebellion are public goods. When a good is categorized as a public good, it means (a) that the good is not able to be excluded from public consumption—that is, the good cannot be privately held—and (b) that the enjoyment of the good by one consumer does not diminish the quality or quantity of the good for another individual. The implication is that an individual who had not aided in the rebellion would still be able to partake of the benefits of a successful rebellion. And given that the costs of a failed rebellion likely would be heavy for participants, both possible outcomes incentivize choosing to stay home instead of joining the group. The emergent paradox is that if everyone were to join, success would be virtually guaranteed, but since the potential costs of joining are so high, few do so at first (Lichbach, 1998). Civil wars occur only when a critical mass of individuals overcomes the incentives to stay home.

To be certain, similar paradoxes arise for other types of collective action, including protesting and rioting. What separates the collective action in civil wars from these other types is that civil wars simultaneously involve collective action (rebellion), political motivations (deposition or preservation of sociopolitical order), and reciprocation (fighting between the rebels and the state). That is, in civil wars, political motivations undergird collective action,

which manifests in violent rebellion to which states violently respond. Civil wars combine collective action with violence that is both political and reciprocal. Each concept is discussed in turn below.

Violence in civil wars is inherently political, which is to say, for instance, that rebels' objectives are based on subverting and changing societies' legal structures. Criminals, by contrast, tend to be less ambitious. Criminal behavior can be described as circumvention of society's legal structures.

The dichotomy between the political and the criminal is critical to defining civil wars conceptually because it permits researchers to scientifically distinguish between the behaviors of the Tamil Tigers and Al Capone's Chicagoland gangsters, for example. The former wanted to live by a sociopolitical order wholly autonomous and distinct from the sociopolitical order imposed by the Sri Lankan central government. Conversely, Al Capone and his ilk sought only to live outside the sociopolitical order and had little desire, if any, to displace or significantly alter the prevailing sociopolitical order.² In a sense, then, antistate behavior wrought by political motivation—as in the Tamils' case—is intentional and systematic. Antistate behavior wrought by criminal motivation is coincidental—a strategy used to navigate around state strictures that occasionally pose inconveniences.

To classify civil wars as inherently political is hardly to isolate them from other phenomena. Indeed, politicide and coups d'état are but a couple of examples of alternative types of political violence. In neither of these types, though, is the violence mutual. In the case of politicide, the state targets civilians who do not and—perhaps—cannot fight back. The converse case applies to coups d'état; essentially, the state is powerless to fight back and subsequently cedes its authority to the usurpers. For the political violence to be associated with civil war, however, there must be *dueling* (explained below) and organized violence between the disputants (Fearon & Laitin, 2003; Sambanis, 2004). For this reason, routs, massacres, or any other type of one-sided violence conceptually differs from civil wars; each can, however, be a precursor to civil war (Kalyvas, 2006; Mason, 2004).

While it is generally accepted that mutual violence between disputants is a necessary condition for a civil war to exist, there is debate about whether the state must be a disputant. The phenomenon of *failed states* in particular has brought this issue to the fore. The modern sociopolitical arrangement requires a state to impose social contracts and to provide its citizens with basic protections from internal and external threats (Migdal, 1988). *Weak states* have limited capacities to do so and often co-opt existing social networks in order to accomplish the most basic operations (Migdal, 1988). In failed states, not only are the most basic state operations neglected, but no coherent apparatus exists from which the state can operate, nor is there a general consensus among the populace of what or, more important, of who the state is (Buzan, 1983). Classic

examples include, among others, Democratic Republic of the Congo, Afghanistan, and Somalia ("The Failed States Index," 2005; Patrick, 2007). The problem for civil war studies is whether the type of fighting typically observed within the territories associated with failed states constitutes civil war, a subtype of civil war, or a separate type of conflict altogether. At issue is the absence of the state as a disputant in the fighting taking place in these countries.

In one of the most important and influential books on conflict, Small and Singer (1982) include among their civil war criteria the requirement that the state must be a disputant in the internal fighting. Yet Sambanis (2004) observes that this principle is not uniformly applied. In Somalia, for instance, a state does not exist in any institutional coherence. There is, consequently, no state involvement in the brutal fighting ongoing within the territory. Still, scholars tend to classify the conflict as a civil war (Sambanis, 2004; see, e.g., Fearon & Laitin, 2003; Sarkees, 2000). To reconcile this incongruence, some degree of nuance has been added.

Foremost, the idea of the state as a disputant has been reinterpreted to include claimants to the state (see Fearon & Laitin, 2003; Sambanis, 2004). A claimant is any sub-state group that asserts its right to fill the void left by, or never occupied by, the state. The fighting in territories of failed states often occurs among multiple, competing claimants (Holsti, 1995). Under this context, such conflicts ostensibly pit disputants against one another for the right to impose their preferred sociopolitical order. As noted above, violent conflict over sociopolitical order within a territory is what makes wars *civil* wars. For this reason, conflicts in Sierra Leone, Afghanistan, and Somalia, for example, tend to be classified as civil wars.³ The state as a disputant is therefore not a necessary requirement for civil war; rather, the disputants must simply be competing over the right to act as the state.

This makes sense considering that this is essentially what states are doing when they instigate or reciprocate violence against coherent substate groups. States fight for the right to impose their sociopolitical order in the given territory, which is exactly what substate groups do in the absence of a coherent state. Consider a few examples.

Secessionist conflicts occur when ethnic enclaves attempt to dissolve their political ties to an existing nation-state in favor of forming their own. Competing claims about who ought to impose whose sociopolitical order therefore evolve from the ethnic enclave and from the extant state. The Chechens' fight to separate Chechnya politically from Russia is an illustrative example of this type of conflict. The Chechens believe they ought to be able to impose their preferred sociopolitical order within Chechnya, but Russia claims the right to project its preferred sociopolitical order from Moscow.

Some intrastate conflicts are believed to emerge because one socioeconomic group perceives itself to be disadvantaged or exploited by another socioeconomic group, which usually controls the state. The fighting that

emerges is a result of the outgroup's desire to depose the existing sociopolitical order, for which the ingroup stands, in favor of a new or significantly altered sociopolitical order. The social revolution in China that took place throughout the 1930s and 1940s represents this type of conflict, as do other rebellions, especially other social revolutions. In that conflict, the Communist Party of China won support from peasants, arguing that the system as run by the Kuomintang exploited the peasants and left them altogether politically powerless. The only way to garner power was to depose the Kuomintang's sociopolitical order and replace it with the Communist Party of China's preferred sociopolitical order. The Kuomintang fought to preserve its right to occupy the state apparatus from which it could impose its preferred sociopolitical order.

Finally, conflicts in territories of failed states, as mentioned above, often involve competing claimants to the state. That is, the substate groups battle for the right to impose their preferred sociopolitical order throughout the territory. Mobutu Sese Seko's departure in 1997 sent Zaire into conflict illustrative of that often found in failed states. Mobutu's departure left a void in central authority. Within a year, violent conflict emerged as Laurent Kabila's Alliance of Democratic Forces for the Liberation of Congo-Zaire, commonly known by the initials AFDL, which changed the name of the country to *Democratic Republic of the Congo*, and Jean-Pierre Bemba's Movement for the Liberation of Congo issued competing claims to the state. Each, in the absence of a coherent state, sought the right to impose his preferred sociopolitical order throughout the Congo.

Although these three examples are by no means exhaustive of the types of intrastate conflict observed, the examples do cover many of the contexts from which intrastate conflict emerges. And despite the differences in the makeup of the disputants in each case, one fundamental aspect of each conflict is striking in its recurrence: The genesis of conflict arises from competing claims to the right to impose the territory's sociopolitical order. The implication is that conceptually, civil wars can be defined without setting firm criteria about the makeup of the disputants. What is more important is the goal of the disputants.

Based on the analysis above, civil wars have been isolated conceptually from other social phenomena. First, civil wars take place within an internationally sovereign territory. Second, civil wars involve collective violence that is reciprocated; one-sided waves of terror do not qualify. Third, the ultimate goal of the disputants is to win or maintain the right to impose their preferred sociopolitical order on the populace residing within a given territory. A working conceptual definition of *civil war* is warranted:

Civil wars are mutually violent conflicts between or among organized disputants who live within the borders of a sovereign member of the international system and who have issued simultaneous, competing claims to the right to impose the sociopolitical order over all, or a segment of, the territory within the system member.

To define civil wars this way jibes with most of the contemporary conceptualizations of civil war by emphasizing the origin and intent of the disputants. Separatist conflicts, social revolutions, rebellions, and conflicts between substate groups trying to fill the void left by the state all fit within this definition. Unfortunately, this definition has flaws, too, which will become evident in the next section.

Measuring Civil Wars

In order to use civil wars as an empirical concept, not only must researchers isolate civil wars conceptually; scholars also must find ways to populate cases from which data can be collected. And for all the progress the literature has made in isolating civil wars conceptually, problems still persist in the operationalization of civil war. For example, the conflict between the Colombian government and the Marxist Revolutionary Armed Forces of Colombia (FARC) fits most of the elements of the civil war definition given in the previous section. Recent developments, though, have led FARC rebels to begin to base their operations out of neighboring Ecuador. The retreat of rebels to neighboring states (Sahleyn, 2007) is but one part of a trend in which civil wars have begun to take on international dimensions (Gleditsch, 2007). Indeed, some fighting between the Colombian government and FARC rebels has taken place in Ecuador. The problem this poses for the present conceptualization is that the FARC rebels no longer live within the recognized borders of the state they seek to depose, nor is all the fighting contained within Colombia, yet the conflict is a widely cited case of domestic insurgency, which suggests the above definition ought to be altered some. But to relax the definition would be to begin to blur the conceptual lines between civil and interstate wars again. This case thus illustrates that endemic problems persist, which makes producing a standard set of civil wars difficult and explains why myriad sets exist. It is therefore perhaps unsurprising that each set tends to favor the idiosyncrasies of its author's research goals (Sambanis, 2004).

Death Totals as Metrics for Civil Wars

Recognizing that problems exist, however, does not imply that the literature is void of any standard approaches to operationalizing civil wars. First, a typical practice is to employ death totals to identify important civil war properties. Total aggregate deaths tend to indicate whether violence up to the level of a war has taken place, and relative death totals among the disputants have been used to indicate whether the violence was one-sided or mutual enough to be considered *dueling* violence (e.g., Fearon & Laitin, 2003; Gleditsch, Strand, Eriksson, Sollenberg, & Wallensteen, 2002; Sambanis, 2004; Sarkees, 2000; Singer & Small, 1994). The proximity of deaths in time also is used to determine the beginning of and, in some cases, the end of civil

wars (see Sambanis, 2004). The differences among the studies include the threshold of deaths necessary and the time over which the deaths must take place. For instance, Gleditsch et al. (2002) use a lower per annum restriction than do Fearon and Laitin (2003) or Sambanis (2004).

The crux of the debate over the utility of deaths as a metric for civil war properties is whether limited skirmishes and overall low levels of violence are sufficiently high to classify as civil wars. If the death threshold is too low, then researchers might include cases in which no *real* threat to the state or extant sociopolitical order exists. Conversely, setting the threshold too high would bias studies toward bloodier conflicts, which may or may not be more divisive than the excluded, less bloody conflicts. The high death total might have something to do with advanced technology, population density, or some other exogenous factor not inherent to the conflict itself. Indeed, Sambanis (2004) discusses the attributes of employing relative measures in order to ensure that the death total substantially affects the nation-state under consideration. For instance, 1,000 deaths in China represent 0.001% of the population, whereas 1,000 deaths would account for about 0.03% of the Uruguayan population and 0.2% of the Montenegro populace. Put another way, deaths in China are worth about 5% of each death in Uruguay and worth about 0.33% of each death in Montenegro. There are also cultural, philosophical, and ideological differences across time and space that place different values on life in general (Finnemore, 2003). For instance, it was just over 200 years ago when the United States constitutionally agreed that people of a certain suppressed and enslaved minority were equivalent to only three fifths of a person from the majority ethnic group. In essence, a literal application of this provision would require nearly 1,700 deaths from the suppressed minority to equate to 1,000 deaths of the majority. For these reasons, deaths are not as straightforward a measure as scholars would like.

The arbitrariness of death thresholds is also an issue. Take Fearon and Laitin's (2003) 1,000-death threshold, for example. Seemingly, there is little substantively different in a civil war from the time of the first death until the time of the 1,000th death. Yet the first death does not become a civil war death until the 1,000th death is reached. Despite these frailties, it is difficult to discern what might be a better, more consistent indicator of political violence than deaths, hence the widespread use of deaths as a metric.

Levels of Analysis

In addition to the issues associated with measuring key properties of civil wars, there is also a debate about what the appropriate level of analysis is for civil war research. The debate prompts researchers to consider whether their measures authentically operationalize their conceptualization of civil wars. Failure to do so could result in an ecological fallacy whereby researchers use one level of analysis to operationalize theories about civil war conceptualized on another level of analysis.

Currently, there are two major camps. One branch of the literature is devoted to resolving macrolevel puzzles about the causes and effects of civil wars (e.g., Collier & Hoeffler, 2004; Fearon & Laitin, 2003; Horowitz, 1985). In effect, the primary question these studies ask is, why do some states experience civil war while others, similarly situated, do not? A second branch of the literature focuses on microlevel puzzles, questioning why some individuals join rebellions when others do not, and what prompts individuals to engage in violence (e.g., Gurr, 1968; Kalyvas, 2003). Some works do a good job of synthesizing the two approaches, showing how macrolevel conditions can affect individuals' choices and, ultimately, their behaviors at the microlevel (e.g., Kalyvas, 2006; Mason, 2004; Mason & Krane, 1989).

The key is that studies must not blur the levels of analysis. Theories about the root causes of civil conflict are no more capable of explaining why individuals commit to mass violence than theories about the desire to punish for past injustices can explain why some states are more prone to civil wars than others. The following examples should make this point clearer. In the former case, consider Horowitz's (1985) seminal work on ethnic conflict. A root cause of civil wars, Horowitz argues, is the need ethnic groups have to compare favorably with each other. When one group compares unfavorably with another, the situation fosters enmity and, eventually, conflict between the groups (Horowitz, 1985). And while this certainly seems to be the case, the theory does not explain the motivation an individual has to commit to violence. That is, not all marginalized individuals join their coethnics to fight, and, moreover, many who do fight do so for reasons that differ from the stated political reasons (Kalyvas, 2003).

Of course, in the latter case, whether individuals are motivated to violence by a sense of justice and retribution or by cold calculation and personal ambition does not explain why some states tend to be more prone to political turmoil than others are. Consider Gurr's (1968) enduring work on relative depravity. Essentially, Gurr (1968) argues that a necessary condition for mass political violence is anger in individuals. The anger evolves from situations when what one thinks one can achieve falls short of what one believes one rightfully should have the opportunity to achieve (Gurr, 1968). That this article is cited frequently four decades after it was written is testimony to its explanatory power. It does, however, fall short of explaining why some states seem more prone to mass political violence than others do, despite the presence of "angry" individuals in each, or why some states produce more relatively deprived individuals than others do. As a microlevel analysis of civil wars, the study cannot explain macrolevel factors.

The point should be made that these are not criticisms of the works inasmuch as they are illustrations of the gap that exists when authors employ different levels of analysis. A few studies do attempt to bridge this gap, however. For one, Mason (2004) discusses the motivations individual

peasants have to join revolts but also demonstrates how certain state properties, such as land-tenure patterns, change peasants' calculi and makes some states more prone to peasant revolt than others are. Mason explains why some peasants choose violence and some do not, as well as why some states seem more prone to peasants who choose violence than other states do. And it is likely that this approach represents the immediate future of the field. That is, integration of the various levels of analysis appears to be on the uptick, although scholars must beware that in integrating, they do not conflate the levels.

Theories About Civil War

Having introduced the most prevalent conceptualizations of civil war, as well as some of the different issues that go along with empirically investigating the concept, this chapter now turns to highlighting some of the most prevalent theories on civil war. This section focuses on theories about the causes of civil war.

To know the root causes of civil wars has been the quest of most civil war researchers. In order to explain the incidence of civil wars, scholars typically take one of two approaches. One approach is rebel based; the other is state based. Theories belonging to the rebel-based approach seek to explain what motivates or encourages substate groups to issue a competing claim to the right to impose a sociopolitical order and, ultimately, to engage in political violence in an effort to substantiate their claim. Two general categories typify the explanations for rebel behavior: greed and grievance (Collier & Hoeffler, 2004).

Greed

Proponents of the greed explanation argue that rebels fight only when there is something to be gained by winning and when the probability of winning is sufficiently high (Collier & Hoeffler, 2004). In other words, civil wars are thought to occur because rebel groups have something economically tangible to gain by winning and have reasonable expectations of winning. The theories within this category therefore tend to focus on variables related to rebels' opportunities.

Two articles demonstrate the key dimensions of the greed, or opportunity, theory. Collier and Hoeffler (2004) demonstrate that rebellions occur when they can be financed. Factors that affect the financing capacity of rebel organizations therefore are critical variables to civil war discussion. One variable is the cost associated with paying recruits. Collier and Hoeffler explain that recruits forgo income to join themselves to a revolution. The benefits they get from joining and fighting must exceed this forgone income. Therefore factors that depress the income recruits could earn outside the rebellion shrink the cost per recruit for rebel organizations, which ostensibly means more recruits and a greater opportunity to succeed (Collier & Hoeffler, 2004).

Fearon and Laitin (2003) show that civil wars are more likely to occur in nation-states where rebels can employ *insurgency*: "Insurgency is a technology of military conflict characterized by small, lightly armed bands practicing guerrilla warfare from rural base areas" (p. 75). Consequently, in employing insurgency, rebel groups reduce their need for more recruits, which drives down the cost of financing the rebellion. The practice of *limited engagements* also reduces costs by lessening the need for stockpiles of ammunition. In a sense, when factors enable insurgency to be used, the likelihood of conflict increases because the costs of financing the rebellion are reduced. So wherever insurgency is a viable technology, rebels have a greater opportunity to finance a rebellion, all else being equal.

The bottom line of each of these two articles is that rebellion is more likely to occur wherever it is more likely to be financed. Sources of financing, therefore, are critical (Collier & Hoeffler, 2004), as are factors that affect the amount of financing required—that is, whether potential recruits must forgo relatively high incomes (Collier & Hoeffler, 2004) or whether rebels can employ a cost-reducing technology such as insurgency (Fearon & Laitin, 2003). To be sure, many more articles touch on other factors related to the greed theory, but these two articles highlight the major facets that link rebels' opportunities to the incidence of civil war.

Grievance

The alternative rebel-based theory of civil war incidence is based on *grievances*. It argues that what drives rebellion is the desire to reform or remake the extant sociopolitical order because of its apparent lack of fairness. According to this theory, rebels fight to rectify social injustices they face. The causes of rebellion, then, are the factors that contribute to perceived social injustices. For this reason, inequality—in a variety of forms—and relative deprivation are the two main foci of the grievances explanation.

Inequality, especially income inequality, aggrieves people (Muller & Seligson, 1987). Inequality by itself, though, does not convert grief to anger and anger to action. Anger is fomented when the inequality is seen as unjust or illegitimate. Such situations occur when the prospects of achieving what one believes one has the right to achieve are low (Gurr, 1968), and such situations are furthermore exacerbated when the relative deprivation is thought to have been foisted on the downtrodden—either formally or informally—by a rival ethnic group (Horowitz, 1985). Under this theory, civil wars represent violent attempts to correct years of perceived mistreatment, not opportunities exploited by rebel entrepreneurs.

Of course, some question whether greed and grievance theories are mutually exclusive. Greed and grievance can interact (Kalyvas, 2003), which would explain why some cases of severe relative deprivation persist and why not all ethnically heterogeneous societies conflict. In the former, despite a reason for rebellion, the opportunity to rebel may

not materialize. In the latter, while resources may be plentiful such that rebellion is financially feasible, there may be no relative deprivation worth rectifying; outcomes might be in line with expectations. In either case, employing one theory seems insufficient to explain the observed behavior. Rebels' motivations in civil wars may at once be born of greed and of grievance.

State-Based Explanations

Much of the literature on why or how rebels fight takes for granted the state's role in civil wars. States tend to be treated as a constant in the civil war calculus. Yet states have many components, and it is not unreasonable to build theories of civil war based on differences in these components. For instance, it could be that certain types of states tend to produce relatively high degrees of inequality or that certain state characteristics tend to yield more and greater opportunities for rebels to finance their operations. Whatever the case may be, there is little doubt that differences in state characteristics exert some effect on the incidence of civil wars. Consider the following examples of state strength and democratization.

Weak states tend to be more prone to civil wars than do strong states (Fearon & Laitin, 2003). The reason civil wars occur in weaker states more readily than in strong states is twofold. First, weaker states tend to be incapable of imposing sociopolitical order throughout their entire territory, which includes providing equal protection and services to all citizens (Migdal, 1988). Weak states tend to be more prone to state-sponsored inequality (Migdal, 1988). This opens space—both figuratively and literally—within which rebels can emerge. Second, weak states are more likely to resort to indiscriminate violence against civilians in order to flush out rebels because the state's intelligence-gathering mechanisms tend to be poorly developed as other aspects of the state (Mason & Krane, 1989). The cumulative effect of state weakness, then, is to give rebels more cause to fight and to make insurgency a more viable option, which in turn makes financing the rebellion more feasible (Fearon & Laitin, 2003). Other theories about state capacity and civil war exist, but most of these are either elaborations or qualifications of the theory as presented here.

The level of democratization also appears to have important implications for states' vulnerabilities to civil war. Democracy preempts rebellion because it decreases the incidence of grievances by giving individuals a voice in the government and permits individuals to coordinate together politically in an effort to work to peacefully remedy perceived inequities (Hegre, Ellingsen, Gates, & Gleditsch, 2001). Coherent autocracies, on the other hand, preempt rebellion by prohibiting the political coordination needed to generate mass violence (Hegre et al., 2001). Where civil war is most likely to occur, then, is among states that suffer from institutional incoherence that fosters grievances, on one hand, and still has elements of openness, on the other hand, which permit individuals to come together and coordinate

their efforts to effect political change (Hegre et al., 2001). Therefore, states characterized by some democratization and some autocratic tendencies tend to be those most prone to conflict. Again, state characteristics are shown to influence the incidence of civil war.

Policy Implications

Understanding the nature of, and causes of, civil wars has profound implications for policy. Presumably, the better scholars understand what factors most likely contribute to civil war, the more likely it is that policies aimed at addressing those factors can be authored and implemented. For instance, consider Horowitz's (1985) discussions about how to resolve ethnic conflict. The resolution directly addresses what Horowitz found to be an important cause of ethnic division, namely, that political fates are often tied to ethnicity. Even democracies can be rife with ethnic conflict if parties form along ethnic lines because again political fate will be directly tied to ethnicity. To resolve this dimension, Horowitz proposes that certain institutions be in place that bring about political parties that cut across rather than reinforce existing ethnic cleavages. That is, parties should be introduced that reflect sociopolitical interests that put coethnics at political odds with each other.

As another example, the fact that institutional incoherence produces vulnerabilities to civil war suggests that researchers should find more effective ways of promoting the transition from autocracy to democracy (Hegre et al., 2001). That is, states appear most susceptible to civil war in the stages of polity transition. The implication is that if democratization is to continue, then ways to minimize the institutional incoherence during transition, as well as to reduce the amount of time spent in transition, are needed.

Finally, understanding the dynamics of the conflicts might also foster policies that can effectively end civil wars that do break out. The statement assumes a link between conflict causes and conflict termination. The veracity of the assumption holds implications for policymakers. In either event, knowing would enable more effective policies to be implemented to end civil wars that do commence.

Conclusion

The scholarly understanding of civil wars is undoubtedly growing. Researchers are now able to wed rebels' motivations and opportunities with theories about state characteristics and civil war onset. For the advancements to continue, however, there needs to be more focus on a standard conceptualization and subsequent operationalization of civil wars. The problem is that researchers will continue to arrive at conclusions based on different cases, which erodes the ability to broadly apply the lessons learned from these studies. Of course, much progress has been made, and civil war research is not the only field that faces

questions about the concept under investigation: Astronomers argue about what makes a planet a planet; physicists still cannot agree on the correct specification of the atom; and literary scholars still debate the identity of Shakespeare. So to say that the concept of civil war needs refinement, then, is not a harsh criticism at all. Rather, it is recognition that better conceptualization and operationalization could lead to discoveries that become significant contributions to the human pursuit of progress and peace.

Notes

1. Moreover, it would seem that the interstate wars in which the systemic order is of issue would be the large, infrequent hegemonic wars where one or more states attempt to establish a hierarchy (Organski & Kugler, 1981).

2. It is perhaps arguable that they even had an incentive to see that the sociopolitical order persisted.

3. It might be reasonable, too, to split civil wars into subtypes because the idiosyncrasies of civil wars in which states are disputants and of civil wars in which claimants are disputants might differ (see Stephens & Liebel, 2008, for example). For the purposes of this chapter, though, it suffices to show that conflicts occurring in territories of failed states can be classified as civil wars.

References and Further Readings

- Buzan, B. (1983). *People, states, and fear: The national security problem in international relations*. Brighton, MA: Harvester Press.
- Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4), 563–595.
- The failed states index. (2005, July 1). *Foreign Policy*, 56–65. Retrieved April 21, 2010, from http://www.foreignpolicy.com/articles/2005/07/01/the_failed_states_index_2005
- Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1), 75–90.
- Finnemore, M. (2003). *The purpose of intervention: Changing beliefs about the use of force*. Ithaca, NY: Cornell University Press.
- Gleditsch, K. S. (2007). Transnational dimensions of civil war. *Journal of Peace Research*, 44(3), 293–309.
- Gleditsch, N. P., Strand, H., Eriksson, M., Sollenberg, M., & Wallensteen, P. (2002). Armed conflict 1946–2001: A new dataset. *Journal of Peace Research*, 39(5), 615–637.
- Gurr, T. R. (1968). Psychological factors in civil violence. *World Politics*, 20(2), 245–278.
- Hegre, H., Ellingsen, T., Gates, S., & Gleditsch, N. P. (2001). Toward a democratic civil peace: Democracy, political change, and civil war, 1816–1992. *American Political Science Review*, 95(1), 33–48.
- Holsti, K. J. (1995). War, peace, and the state of the state. *International Political Science Review/Revue internationale de science politique*, 16(4), 319–339.
- Holsti, K. J. (1996). *The state, war, and the state of war*. Cambridge, UK: Cambridge University Press.
- Horowitz, D. (1985). *Ethnic groups in conflict*. Berkeley: University of California Press.
- Humphreys, M., & Weinstein, J. (2008). Who fights? The determinants of participation in civil war. *American Journal of Political Science*, 52(2), 436–455.
- Kalyvas, S. (2003). The ontology of political violence: Action and identity in civil war. *Perspectives on Politics*, 1, 475–494.
- Kalyvas, S. (2006). *The logic of violence in civil war*. Cambridge, UK: Cambridge University Press.
- Lichbach, M. I. (1998). *The rebel's dilemma*. Ann Arbor: University of Michigan Press.
- Mason, T. D. (2004). *Caught in the crossfire: Revolutions, repression, and the rational peasant*. Lanham, MD: Rowman & Littlefield.
- Mason, T. D., & Krane, D. A. (1989). The political economy of death squads: Toward a theory of the impact of state sanctioned terror. *International Studies Quarterly*, 33(June), 175–198.
- Migdal, J. (1988). *Strong societies and weak states: State society relations and state capacities in the third world*. Princeton, NJ: Princeton University Press.
- Muller, E. N., & Seligson, M. A. (1987). Inequality and insurgency. *American Political Science Review*, 81(2), 425–451.
- Olson, M. (1971). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Organski, A., & Kugler, J. (1981). *The war ledger*. Chicago: University of Chicago Press.
- Patrick, S. (2007). “Failed” states and global security: Empirical questions and policy dilemmas. *International Studies Review*, 9(4), 644–662.
- Sahleyn, I. (2007). Transnational rebels: Neighboring states as sanctuary for rebel movements. *World Politics*, 59(2), 217–242.
- Salter, M. B. (2002). *Barbarians and civilization in international relations*. London: Pluto.
- Sambanis, N. (2004). What is civil war? Conceptual and empirical complexities of an operational definition. *Journal of Conflict Resolution*, 48(6), 814–858.
- Sarkees, M. R. (2000). The correlates of war data on war: An update to 1997. *Conflict Management & Peace Science*, 18(1), 123–133.
- Singer, J. D., & Small, M. (1994). *Correlates of war project: International and civil war data, 1945–1992*. Retrieved April 21, 2010, from <http://www.icpsr.umich.edu/cocoon/DDI/SAMPLES/09905.xml>
- Small, M., & Singer, J. D. (1982). *Resort to arms: International and civil war, 1816–1980*. Beverly Hills, CA: Sage.
- Stephens, M. I., & Liebel, S. R. (2008). *Civil war, anarchy, and conflict management*. Paper presented at the annual meeting of the Midwest Political Science Association, Chicago.
- Vazquez, J. A. (1993). *The war puzzle*. Cambridge, UK: Cambridge University Press.
- Waltz, K. (1959). *Man, the state, and war: A theoretical analysis*. New York: Columbia University Press.

14

TERRORISM

ERIK S. CASE

University of North Texas

Since September 11, 2001, considerable attention has been devoted to the study of terrorism, yet scholarly analysis of the subject has actually been active for several decades. With this increased focus, confusion has arisen as to the very meaning of *terrorism*. In addition, there are competing theories in regard to the causes and effects of terrorism, with contributions coming from economists, sociologists, psychologists, and political scientists. The study of terrorism is truly a multidisciplinary endeavor. This chapter provides a review of the debate regarding the definition of terrorism, presents historical examples of terrorism to provide context, and introduces the primary theoretical and empirical contributions of major scholars in the field.

Defining Terrorism: A Distinct Form of Political Violence?

The term *terrorism*, like *globalism*, is difficult to define and has a diversity of meanings among different groups and individuals. As a common cliché says, “One man’s terrorist is another man’s freedom fighter.” The shifting contexts in which the term is used make it difficult, but not impossible, to study the phenomenon as a distinct form of political violence. For the purposes of empirical analysis, terrorism must be defined explicitly. This chapter offers such a definition, while acknowledging that it

may differ from that of other scholars, cultures, governments, media outlets, and perhaps the reader. It is useful to examine first the evolution of the usage of the term throughout history. Although examples of terrorism stretch back several millennia, the word *terrorism* is relatively new to the world stage.

A Historical Review of the Terminology

The first usage of “terrorism” was in reference to the actions of a nation, not a subnational group. After the French Revolution in the late 18th century, the victors conducted a brutal campaign against nobles perceived to be enemies of the newly formed state. The series of mass executions by the postrevolutionary government was referred to as *the terror*. This early conceptualization differs from the more modern use of terrorism, in which the perpetrators are not usually governments and are instead nonstate actors (Laqueur, 2001). The usage of *terrorism* to refer to the violent actions of nonstate actors arose in response to the bombings and assassinations conducted by radical members of political movements such as anarchism and revolutionary socialism in the late 1800s and early 1900s. During the 20th century, the use of the term expanded immensely, often to include any type of political violence that the observer found to be disagreeable. This is where the murkiness of *terrorism*’s meaning arises. Politicians and the media are quick to label any enemy violence as

terrorism. When any act of violence one disagrees with constitutes terrorism, the concept loses its meaning as it has become highly contextual and subjective. This is why we see certain acts of violence covered as terrorism by certain media outlets and as legitimate resistance by others. The atmosphere of confusion is not helped by the fact that, between governments and researchers, there are more than 100 different working definitions of terrorism. In addition to these observations, Jenkins (1974) pointed out the relativistic nature of the term when he wrote that terrorism seems to mean simply whatever the “bad” guys are doing. Merari (1993) echoed this when he noted that the term had become more of a derogatory epithet than an adjective describing a unique phenomenon.

Developing an Explicit Definition

If *terrorism* is simply a subjective catchall term for many types of political violence, why then is it given its own chapter in this handbook? The answer is that, although the expression can be carelessly or intentionally misused by political figures and media outlets, there is a general consensus among scholarly researchers that terrorism is a distinct form of violence, different from riots, coups, inter- and intrastate warfare, and so forth. Tilly (2004) contended that the definition should be based on characteristics of perpetrators and victims. That is certainly a good starting point.

First, an examination of the nature of the victims is in order. By far, this is the most controversial definitional aspect of terrorism. Take, for example, the 1983 suicide bombing of a United States marine barracks in Lebanon by the militant group Hezbollah. The strike occurred after U.S. troops were sent in to mediate an increasingly violent civil war, and well over 200 soldiers were killed. The attack is almost universally referred to as terrorism, but this is problematic. The victims in this case were armed security personnel stationed in a war zone. If every surprise attack on active armed forces is considered terrorism, there is little to distinguish terrorism from unexpected attacks that occur in conventional warfare. Hence, many scholars believe that terrorism involves violence directed at civilians. That is to say, the victims of terrorism are not actively or officially involved in a violent conflict.

Second, it is important to distinguish the characteristics of the perpetrators of terrorism. History is replete with examples of nation-states targeting civilians, whether their own citizens or individuals in other countries. Traditionally, though, such actions are referred to as *state terror* or *war crimes*. More often, those designated as terrorists are members of subnational groups, meaning they do not have the characteristics of a modern state, such as holding a monopoly on the *legitimate* use of violence in keeping with the classic Weberian definition in an internationally recognized, geographically defined territory or fielding a conventionally equipped army and navy. Examples of such nonstate actors include transnational groups such as

al Qaeda or regionally based rebel groups such as the Basque separatists Euskadi Ta Askatasuna (ETA) in Spain. It follows, then, that another criterion for a violent act to be considered terrorism is that it be committed by nonstate actors. Note that excluding state terror or war crimes from falling under the rubric of terrorism does not pass a moral judgment regarding the reprehensibility of such acts. In fact, the goal of this section of the chapter is to arrive at a definition that is as free from subjective moral, or “normative,” conditionality as possible. Discarding some of the emotional and moral baggage that accompanies the loaded term will allow us to arrive at a more analytically useful conceptualization.

Beyond the nature of the attacker and the victims, it is important to analyze motivations behind terrorist acts. If we were to stop with the definition we have now, any violent crime committed by one civilian against another would be considered terrorism. However, there appears to be something qualitatively different that distinguishes the violent crimes committed by individuals and groups motivated by the desire for economic gain from the violent crimes committed by terrorist organizations. For example, there appears to be a fundamental difference between violent groups such as Hamas in the Palestinian territories and violent groups such as the Sicilian Mafia in southern Italy. Both are nonstate actors that use violence against members of a civilian population, but the distinction lies in the goals of the organizations. Hamas desires to destroy the Israeli state and claims to fight for a Palestinian homeland—a political goal. Organized crime outfits such as the Mafia have no such political intentions. They may have political connections in the form of corruption and bribery, but their aspirations essentially revolve around making money. Groups that use terrorism have political goals that extend beyond the immediate consequences of an attack. In other words, they seek to effect political change by committing violent acts against civilians. Political change can take a variety of forms. For example, terrorism can be used with the aim of changing or gaining policy concessions from a particular government or of destabilizing that government altogether. It can also be used to gain domestic support from an aggrieved group by demonstrating a willingness to fight for a cause. Thus, a final condition for a violent act to qualify as terrorism is that it be motivated by a desire to influence a political outcome. This is not to say that those who use terrorism cannot also engage in organized crime (the Taliban in Afghanistan and the Fuerzas armadas revolucionarias de Colombia, [Revolutionary Armed Forces of Colombia; FARC] turn huge profits by trafficking in narcotics), but their goals are primarily political in nature. This last condition excludes several other types of violence. For example, hate crimes without any known motivation beyond the immediate act would not be considered terrorism. Violent crimes committed by the mentally disabled, such as the attempted assassination of President Ronald Reagan by John Hinckley, also do not have an ultimate political motive and would not fall under terrorism.

Therefore, the working definition of terrorism developed in this chapter is as follows:

Terrorism is the use of violence against civilians and is perpetrated by nonstate actors with the intent of achieving some political outcome.

This definition is by no means the authoritative final word, and there remain several gray areas. Some terrorist groups have skillfully stretched the meaning of *combatants* to apply to virtually every citizen of the state that they oppose. Below are some alternate definitions illustrating the fact that, even within the U.S. government, competing definitions exist:

Terrorism is the use of force or violence against persons or property in violation of the criminal laws of the United States for purposes of intimidation, coercion, or ransom. (U.S. Federal Emergency Management Agency, 2009, para.1)

[Terrorism is] premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents, usually intended to influence an audience. (U.S. State Department, 2001, "Definitions," para.1)

Terrorism can be practiced alongside other modes of violence. For example, a rebel group can choose to use hit-and-run attacks against military targets while also conducting strikes against a civilian population. In such a case, terrorism is one component of the group's overall strategy of warfare. It can be expanded, put on hold, or altered according to the decisions of the group's leader or leaders. This occurred in the case of the African National Congress (ANC), which used terrorism in its fight against the South African apartheid-era government. When the apartheid government finally collapsed and the ANC took over the reins of power, it no longer conducted terrorist attacks. As a result of such examples, terrorism is treated in this chapter as a *tactic that can be used strategically*. Repugnant though it may be, terrorism is a method for achieving an objective, and *terrorist groups* may more accurately be considered to be groups that use the tactic of terrorism.

Historical Examples of Terrorism

The works of Enders and Sandler (2006), Laqueur (2001), Hoffman (1998), and White (1998) together provide a fascinating account of the history of terrorism. Although the term is relatively new, there are many historical examples of nonstate actors using violence against civilians with the broader goal of influencing a political outcome. One of the famous early examples was that of the Sicarii. The group's name derives from its preference for using *sica*, or long knives, when attacking its targets. During the middle of the 1st century, Palestine was occupied by the Roman Empire. A rebellious Jewish group known as Zealots chafed under Roman rule and actively agitated for the removal of the

Roman regime. Within the Zealots, a more radical sect existed, the Sicarii, who felt that violent overthrow was the most appropriate course of action. The strategy of the Sicarii was what would become a classic element of terrorist violence: to provoke a government (in this case, that of Rome) into an overreaction that would, in turn, drive more supporters (in this case, fellow Jews in Palestine) toward the terrorists' cause. The Sicarii embarked on a campaign of assassinations directed against Roman officials and fellow Jews deemed to be collaborators. Historical reports indicate that the Sicarii assassinated their victims in broad daylight and among crowds of people to maximize the dramatic effect. The group also attacked infrastructure, damaging the water supply and destroying agricultural and financial targets.

Operating from the 11th to the 13th centuries in Persia and Syria, a group known as the *Assassins* was a violent splinter group from the Ismaili religious sect. The Assassins were a small group motivated by a desire to protect their religious practices from repression by rival factions, and they held sacred the act of eliminating their victims with daggers. Because of the limited manpower of the Assassins, their leader realized they could not confront the government head-on and instead opted for a sustained campaign of assassinations. Among their most prominent victims were high officials of government, including the king of Jerusalem. The premium that the Assassins placed on dying for their cause, which they believed to be an act of martyrdom, would be echoed by religious extremists in far more recent times.

Religious fanaticism and radical separatism are not the only motivating ideology among groups that used terrorism in history. In the mid-19th and early-20th centuries, a political movement known as *anarchism* became popular in Europe and North America. The ideology advocated the dissolution of all forms of government, deeming them inherently exploitative and unjust. Some anarchists rejected the more traditional means of political expression, such as protests and the dissemination of propaganda through pamphlets. Rather, they advocated what has been called "propaganda by deed." Using violent action to demonstrate by example is a thematic element found among practitioners of terrorism. Violent anarchists in France, Italy, Britain, and the United States were not coherently organized groups, and their targets and methods differed. In Italy and France, there were several high-profile assassinations of government officials and merchants deemed to be part of the corrupt state apparatus. The technological advances that have occurred since the Assassins now allow for a new type of terrorism that has become emblematic of the modern era: the use of bombs. Dynamite was a weapon of choice among violent French and American anarchists, and the use of explosives in terror attacks has been adopted by an ever-widening array of groups in the last two centuries.

The systematic campaign of violence and intimidation directed toward Blacks and moderate Whites in the United

States by the Ku Klux Klan (KKK) is an infamous example of domestic, “homegrown” terrorism. The KKK was devoted to preserving the traditional dominance of Whites in the U.S. South, and this vision entailed violent actions designed to forcibly dissuade Blacks from participating in the political process. A ritualistic right-wing organization, the KKK was indelibly known for its practice of wearing white hoods and its pseudoreligious practice of burning crosses. KKK activities included a series of extrajudicial (outside the law) executions known as *lynchings*. The KKK used lynchings and other forms of violence to murder numerous individuals throughout the 19th and 20th centuries. In the 1960s, the KKK caused an international uproar when it planted a bomb in the basement of the 16th Street Baptist Church in Birmingham, Alabama. The resultant explosion killed four young girls and served to galvanize the civil rights movement that was emerging in the United States (National Public Radio, 2003).

Terrorism was a central strategy for gaining power within several political movements during the 20th century. Communist revolutionaries were among several groups using terrorism to agitate against czarist rule in Russia in the 1900s. By the time they achieved power in a 1917 coup, they had conducted three major campaigns of assassination against political enemies. The fascist Nazi party also used terrorism in its ultimately successful attempt to gain control over the German government prior to World War II. Thus, two of the most historically significant political movements of the 20th century, fascism and communism, made use of terrorism to successfully transition from being challengers of governments to becoming the governments themselves.

The post-World War II era heralded several new developments that came to characterize modern terrorism. First, the growing popularity and availability of commercial air flight presented new opportunities for groups aiming to conduct terrorist attacks. During the 1960s and 1970s, *skyjacking* was used effectively and often. In these hostage-taking situations, aircrafts were forcibly commandeered in flight or on the ground, and political demands were made by the hijackers. Defensive countermeasures taken, such as the installation of metal detectors at airports, helped to greatly reduce the number of terrorist hijackings in the 1980s and 1990s. Second, the 1980s saw the advent of suicide bombing, first used by the Lebanese group Hezbollah and refined by the Liberation Tigers of Tamil Elam (LTTE) separatists in their fight for an autonomous homeland separate from the Sri Lankan government.

Al Qaeda and the September 11 Attacks

Following the September 11, 2001, attacks, the U.S. government tasked the National Commission on Terrorist Attacks Upon the United States to create a comprehensive account. The Commission produced the *Complete 9/11 Commission Report* (2004), which is a massive and freely

available online resource that is easy to read and highly recommended for students interested in a more thorough accounting of the attacks. The following account is primarily derived from the *Complete 9/11 Commission Report*.

In 1978, the Soviet Union invaded the Central Asian nation of Afghanistan. A rugged geographic region characterized by a tribal society, Afghanistan was home to several million Muslims of differing ethnic backgrounds. The possibility of an atheistic Communist regime occupying a Muslim nation proved to be an incendiary prospect for many Muslims across the world. It also presented an opportunity for governments in the Middle East, such as Saudi Arabia and Pakistan, to export young radicals to fight the Soviet occupation. In doing so, these governments gained stature as defenders of the Muslim faith while also relieving themselves of the internal security threat these radicals represented. The United States, Soviet Russia’s primary cold war enemy, also provided funding for some of the groups battling the Soviets. These groups were by no means uniform in their methods and motivations, and after their victory over the Soviets in 1989, they proceeded to battle among themselves, plunging Afghanistan into another decade of statelessness and civil war. One of the many foreigners who traveled to Afghanistan to fight against the Soviets was Osama bin Laden, the son of an immensely wealthy Saudi Arabian construction magnate. After the Soviets left Afghanistan, bin Laden was instrumental in maintaining a cadre of followers that eventually became known as al Qaeda.

Bin Laden was largely the financier of al Qaeda, and his deputy, Egyptian physician Ayman al-Zawahiri, headed up operational control of the organization. Espousing a radical version of Sunni Islam and inspired by the Muslim Brotherhood, the al Qaeda leadership was determined to battle any party it perceived to be hostile toward its austere version of Islam. Bin Laden blamed the suffering of fellow Muslims on corrupt Middle Eastern regimes, whom he perceived to have betrayed the faith, and the actions of Western governments, particularly the United States. When Iraq invaded Kuwait in 1991, the neighboring Saudi regime became concerned that it was Saddam Hussein’s next target. Bin Laden offered the protective services of his fighters, but the Saudis, wary of importing the same figures that helped topple the Soviets, instead opted for protection from the United States, allowing U.S. military forces to use Saudi territory as a launching pad for the first Gulf War. The prospect of foreign troops on Saudi soil (home to the two most holy sites in Islam: Mecca and Medina) proved to be intolerable to al Qaeda. The organization also held grievances against the United States for the country’s support of Israel and other Middle Eastern governments that al Qaeda held a deep enmity toward. Thus, the United States became a primary target of the organization.

In 1998, bin Laden and al-Zawahiri held a news conference to announce the issuance of a religious edict. In it, they claimed it was the duty of every observant Muslim to attack Americans whenever and wherever possible. Soon

afterward, bin Laden clarified that the group made no distinction between American troops and civilians, reasoning that civilians in a democratic society were directly responsible for the policies of a government they voted for. Several terrorist attacks attributable to al Qaeda had already occurred by the time the 1998 edict was issued, but none at that point had been as large as the two simultaneous vehicular bombings of U.S. embassies in Sudan and Kenya that followed that summer. The bombings killed hundreds of civilians, mostly Muslim East Africans. This irony did not go unnoticed by al Qaeda, which, hoping to be the vanguard for disaffected Muslims throughout the world, had already embarked on more ambitious plans to attack Americans directly.

Some time that same year, bin Laden purportedly gave approval for the September 11, 2001, attacks. The operational planning for this task was largely delegated to a man known as Khalid Sheikh Mohammed. Born in Kuwait and college educated in the United States, Khalid Sheikh Mohammed reportedly claimed to have gotten the idea for a plane attack on the World Trade Center (WTC) after his nephew's 1993 attempt to destroy the towers with a vehicle bomb failed. Throughout 1999, the plan progressed as candidates for the operation were recruited, trained in Afghanistan, and secured visas for entry into the United States. While in the United States in 2000 and 2001, the hijackers attempted to blend in as visiting students, and select members attended flight training school.

By 2001, the U.S. government had indications that al Qaeda was planning a massive attack on the homeland, but the government was unable to connect the bits and pieces of intelligence, spread across several governmental agencies, into a coherent enough picture to stop the plot before its execution. The nature of the attack was also unexpected because there had never been one like it before. There had been suicide bombings in other countries, but no one had ever combined suicide attacks with airline hijacking. The conventional model of hijacking and hostage negotiation was not part of the plan, and few, if any, predicted that entire commercial airliners could be taken over and commandeered with items as seemingly innocuous as box cutters. Thus, most of the world was shocked when 19 men hijacked four planes on September 11, crashing three of the planes into their intended targets—the WTC towers and the Pentagon—while the fourth was brought down in a Pennsylvania field when the passengers fought to retake control of the aircraft. The attacks resulted in the destruction of the WTC towers and part of the Pentagon, with more than 3,000 civilians murdered. The size and scope of the operation are without parallel.

Historical Patterns

A few patterns emerge in this brief, and by no means exhaustive, historical review of the occurrence of terrorism. These patterns have guided much of the research in

academia. First, groups that employ terrorism come from a variety of ideological backgrounds. Some, like the Assassins, were religiously motivated. Others were devoted to a left or right wing political cause, such as the anarchists or the KKK. Groups such as the LTTE aimed to carve out a piece of territory for their ethnic kin. These motivations are not mutually exclusive. For example, the Sicarii were a religiously motivated group that *also* sought an element of territorial control. The primary lesson to draw is that no one extreme political or religious ideology dominates the use of terrorism. The backgrounds and motivations of these organizations are highly diverse, and attempting to uncover a single, overarching ideology among them all is a fruitless endeavor. Second, a similar amount of diversity is found across time and geographic region. Although the methods of implementation have shifted with technology, terrorism is not a new phenomenon, and it is not restricted to one region of the world. Third, all the previously mentioned groups, at least initially, lacked the size and ability to field regular armies that could mount a direct challenge to governmental authorities. This makes sense because, by definition, groups that use the tactic of terrorism are nonstate actors. However, some movements have grown large enough to challenge authorities by more conventional military means, being able to simultaneously pursue guerrilla campaigns or fight traditional battles while also conducting terrorist strikes. Last, the attacks conducted by various terrorists were intended to provoke a response, whether from the element of society they were challenging, the societal groups they claimed to be fighting for, or both.

Theoretical and Empirical Contributions to the Study of Terrorism

Terrorism studies, as a subfield of political science, are a relatively new endeavor. They do not have the level of development we find in other arenas of political science, such as the study of interstate war or internal revolutions. Despite that, there have been some major developments in the field during the past three decades. This portion of the chapter analyzes some of the seminal works on terrorism and reports some of the critical empirical conclusions from this area of study.

The Causes of Terrorism

Individual and Group Motivations

As is often the case with the study of some unique social phenomenon, the first efforts were primarily aimed at explaining its causes. There is also much to be learned from the effects of terrorism, but understanding the origins of terror is of primary importance. Jenkins (1974) produced an analysis for the U.S. Congress that contained considerable insight into the nature of terrorism. Jenkins

forcefully argued against the perception that terrorism was the work of senseless, mindless, and irrational actors. Attacks on civilian targets shock the conscience, and the seemingly random, chaotic carnage that is produced by terrorism understandably gives the impression to many that it is the work of insane individuals. This impression may even be intentionally cultivated by terrorist practitioners themselves, and the media also play a role in casting terrorists as lunatics, but the evidence belies this assumption. Jenkins's work noted that practitioners of terrorism had concrete political goals and did not simply engage in violence for the sake of violence in the way that sociopathic criminals might. Rather, because terrorist groups often lack the resources to mount a direct challenge to government security forces, they shift their focus to "softer," less fortified targets, such as civilians. The often indiscriminant nature of attacks on civilians garners the most attention, so while the purpose of a more conventional military operation may be to take and hold a piece of territory, the terrorist attack is designed to influence an audience beyond that of the immediate victims. Often the goal is to create a climate of panic and to expose a government's inability to prevent such indiscriminate violence. Because of these unconventional goals, Jenkins cast terrorism as a form of political "theater" in which the harm suffered by the immediate victims of terrorism is of secondary importance to the group conducting the attack. Take, for example, the September 11, 2001, attacks. The perpetrators chose targets that were steeped in symbolism. The Pentagon and the WTC were prominent symbols of American financial and military might, readily recognizable as such by both American citizens and the worldwide audience that witnessed the attacks. The attacks sent a message that the citizens and institutions within the United States mainland were not safe, and the atmosphere of fear following the attacks was palpable. The name Osama bin Laden gained universal recognition, while al Qaeda became the object of unceasing media attention. Yet despite the disastrously large death toll, the average American citizen's chance of being harmed in such an attack remained infinitesimally small. For the average person, the chance of being felled by a heart attack or car accident was far greater than that of being struck down by al Qaeda, yet terrorism became the central issue for several election cycles following the attack. These facts give credence to Jenkins's argument that terrorism is theater: a spectacle designed to attract maximum attention and create massive emotional impact. It became clear that one of the reasons terrorism occurs is the massive amount of attention it can attract without being cost prohibitive for small organizations.

Crenshaw (1981) was one of the earliest political scientists to conduct research on terrorism. She concluded that terrorism was not necessarily the result of broad public dissatisfaction with the political order or evidence of a fractious society. Rather, Crenshaw contended, terrorism was often the result of the grievances of a disaffected group that had originated in the elite and claimed to fight

for a larger group. This conclusion was reasonable given the makeup of left-wing terrorist groups that had dominated the news in the previous two decades. While claiming to fight for the downtrodden worker, groups such as the Red Army Faction in Germany were largely composed of students from upper- or middle-class origins. Their parents were academics, clergy, writers, and other professionals, yet the students became disaffected and alienated from the societies that spawned them. Crenshaw believed that psychological factors such as guilt, desire for vengeance, and a thirst for excitement were the primary motivations of individuals who participated in terrorism.

The idea of a *rational terrorist* has very real policy implications for counterterrorism officials. If an enemy is mentally deranged, irrationally lashing out at random targets, there is little use in trying to predict when, where, and how that enemy will strike. However, if an enemy is calculating, weighing the costs and benefits of conducting an attack, policymakers and analysts are more likely to be able to get an idea of what targets the enemy will select and how it will attack them. It is no surprise, then, that many researchers take the rational choice approach when modeling terrorism. In this research strategy, the choice of whether to participate in terrorism is contingent on the cost-benefit ratio for the principal actor, the terrorist. Scenarios are often presented as a game wherein an actor takes a turn, choosing whether to use terrorism, and another actor, usually representing a government or counterterrorist agency, must choose a response while considering that the terrorist opponent will be trying to predict the agency's strategy (Lake, 2002). The key, then, is determining what particular conditions create a payoff structure for potential terrorists sufficient to entice them to commit the violence despite knowing what the counterresponse will be. More simply put: What conditions cause terrorism to be worthwhile to certain groups and individuals despite the risk? The following section describes various theories regarding those conditions. Attributing thought and rationality to terrorism is controversial as many are reluctant to admit any quality but insanity to such a reprehensible form of violence, but the idea of the *strategic terrorist* is one of the dominant themes in academic research on the matter.

However, the rational choice approach has been criticized as unrealistic. Individual interviews of terrorists have often revealed no such cost-benefit analysis among the terrorists that were studied. Sometimes the choice to participate is based on a desire for revenge rather than a sober calculation of the possible costs and payoffs. One failed suicide bomber noted his spiritual motivations for engaging in terrorism, an inspiration that is not easily modeled by rational choice (Hassan, 2001). Perhaps the only blanket statement that can be made about the "average terrorist" is that he or she is willing to undertake extreme risk. That said, different motivating factors may drive the foot soldiers of terrorist groups and their leaders. It seems more reasonable to expect that the leaders of groups that utilize

terrorism do look at the “big picture” and consider the benefits and risks of various oppositional strategies.

Structural Causes of Terrorism

Rather than focus on the individual psychological calculations of the individual terrorist, some researchers have put forth causal arguments based on the institutional and structural features of a society. One such argument revolves around the supposed connection between poverty and terrorism.

In both the literature and the culture at large, there is an expectation of a causal relationship between poverty and terrorism. Following September 11, 2001, politicians such as former Vice President Al Gore and President George W. Bush argued that combating terrorism should involve efforts to eradicate poverty and increase education in the world’s troubled hot spots. Academics too have prescribed the lifting of living standards in various regions in the hope of creating a disincentive for participation in terrorist activities. There is good reason to think that certain socioeconomic factors are determinants of terrorism. The lack of economic and educational opportunities has already been empirically linked to a variety of other problems, such as property crime, the occurrence of civil war, and instability within new democracies. The general theory is that poverty and a lack of opportunity increase the level of grievances among economically marginalized members of society and that a subset of an aggrieved population may choose to express its discontent violently by way of terrorism. However, the actual evidence of a connection between poverty and terrorism is mixed at best. Counterfactual examples include of the finding that many of the perpetrators of the September 11, 2001, suicide attacks came from educated, middle-class backgrounds. A profile of failed Palestinian suicide bombers in *The New Yorker* returned the same results (Hassan, 2001). Most of the young men interviewed had held jobs, were educated, and did not come from extreme poverty. On the other side of the Israeli–Palestinian conflict, an analysis of the Israeli Jewish Underground, an organization that attacked Palestinian civilians during the 1970s and 1980s, also found that a strong majority of members were highly educated and held prominent occupational positions. There is the possibility that although poverty or lack of educational opportunities is not prevalent among the practitioners of terrorism, it may be that poor socioeconomic conditions experienced by their ethnic or religious kin inspire a so-called Robin Hood model of terrorist activity. If this were the case, it would be more appropriate to analyze aggregated societal or country-level indicators of economic and educational conditions than the individual socioeconomic origins of the terrorists themselves. The empirical evidence does not conform to the conventional wisdom on that level of analysis, either. For example, a survey in the Palestinian territories found that unemployment actually reduced support for terrorism against the Israelis. Another study found

that, after controlling for civil liberties (on account of the fact that poorer countries are more likely to have fewer political freedoms), there is no statistically significant difference between poor and rich countries in terms of the number of terrorists that they spawn (Krueger & Maleckova, 2003).

The Effects of Terrorism

The rational choice approach is particularly helpful in examining the strategic expectations that terrorist practitioners have for the consequences of their acts. Specifically, there is evidence that terrorism is used to goad an overreaction from the target government or societal group (Bueno de Mesquita & Dickson, 2007; Lake, 2002). This strategy appeared to be in play among jihadist elements of the Sunni insurgency in Iraq.

Following the 2003 invasion by the United States, Iraq experienced an influx of religiously motivated fighters from neighboring Muslim countries (quite similar to the influx of fighters into Afghanistan in response to the Soviet invasion), who claimed to battle on behalf of the minority Sunni population. The Sunnis had held a position of privilege over their Shi’a counterparts during the rule of Saddam Hussein and were now facing the possibility of losing that status. These fighters, along with indigenous radical Iraqi Sunnis, formed the backbone of what would be known as al Qaeda of Mesopotamia. This regional al Qaeda franchise embarked on a series of gruesome executions and suicide bombings directed at the majority Shi’a population, killing scores of civilians. While some Shi’a leaders counseled restraint in the face of such attacks, the Shi’a population evinced a growing inclination toward self-protection and revenge. By the time al Qaeda of Mesopotamia blew up the Golden Mosque, one of Shi’a Islam’s holiest shrines, reprisal killings directed at Sunni civilians were well under way by Shi’a militants. In the following months, the country of Iraq descended into a vicious period of civil conflict, with spasms of indiscriminate killing on both sides (Frontline, 2007).

There is a coherent logic in attempting to spark a wider conflict through the use of terrorism. In provoking a disproportionate response by using terrorism, radical groups are able to pressure or shame moderate members of their community into supporting a violent approach. An extremist group without a large base of support can use as a recruitment tool the overreaction that they have intentionally caused. The Iraqi example was largely one of a disproportionate response from nonstate actors from one element of society directed at nonstate actors from another, but there are also examples of groups using terrorism to cause a government to overreact. The end result is the same: A disproportionate government response can radicalize previously fence-sitting moderates and drive them to support the extremists in their midst. Several rational choice studies have explicitly identified the causal link between overreaction and its effect of provoking further

violence (Bueno de Mesquita & Dickson, 2007; Mason & Krane, 1989). Specifically, governments and subnational groups seeking to respond to a terrorist attack often suffer from an information problem. Because of the secretive nature of terrorist organizations, it is exceedingly difficult to identify the specific perpetrators and punish *only* them. Often, the counterattack response to terrorism is indiscriminate, harming people who had no connection with the original terrorist activity. This was certainly the case in Iraq, where simply having a Sunni-sounding last name was often enough to warrant execution by Shi'a militia members. Imagine, then, how the calculated expectation of the risk of participating in violence changes for an individual in the face of indiscriminate violence. Beforehand, the level of risk derived from joining a violent organization exceeded the potential payoff gained by simply sitting on the sidelines and waiting the conflict out. In the face of indiscriminate violence, where one is equally likely to be harmed regardless of whether he or she had previously participated in violent activity, it becomes less risky to join a violent organization. In fact, it may appear to be in an individual's *best* interest to join such a group because at least the individual is offered an element of protection by doing so. It is no surprise, then, that from the Sicarii to al Qaeda in Mesopotamia, terrorism has been strategically employed to foment higher levels of violence and gain supporters.

The Economic Effects of Terrorism

Assessing the economic impact of terrorism has proven to be one of the most quantitatively rigorous areas in the field of terrorism studies. Economic damage has been an implicit or explicit motive behind several terrorist movements. Osama bin Laden drew lessons from the economic toll inflicted on the Soviet Union during its invasion and occupation of Afghanistan and advised his followers to conduct strikes on Middle Eastern oil facilities in hopes of causing similar economic pain on the energy-reliant Western nations (Associated Press, 2004). Based on the logic that violence will reduce an important source of revenue for the target government, many high-profile attacks are conducted at popular tourist destinations. Indeed, an economic analysis by Enders and Sandler (1991) of terrorist attacks in Spain found that the average attack caused the number of tourists visiting the nation to decrease by approximately 140,000 people in a year. A similar investigation of terror attacks in Italy determined that a typical strike resulted in a shock to the level of tourism that took a full year to dissipate (Enders, Sandler, & Parise, 1992). Beyond affecting specific industries, terrorism appears to have macroeconomic consequences as well. It is interesting that, although terrorist attacks appear to have only a temporary, and small, negative impact on the gross domestic product of nations, a more significant and positive relationship exists between terrorism and government spending (Blomberg, Hess, & Orphanides, 2004). While the goal of

an attack may be to cause economic harm by bleeding a government, an ironic consequence is that governments respond to an attack by investing in counterterrorist measures that require matériel and personnel, thus mitigating some of the economic damage caused by the strike.

The Consequences of Suicide Attacks

Scholarly analysis has demonstrated that suicide bombing, although horrific and repugnant, can produce tangible benefits for groups seeking political change through violence. Sprinzak (2000) noted that suicide bombing is one of the most psychologically effective methods because it communicates the message that there is no deterrent that can dissuade the attacker. Before the popularization of suicide bombing by groups such as Hezbollah in Lebanon and the Tamil Tigers in Sri Lanka during the 1980s, it was largely, and erroneously, assumed that, although willing to engage in extreme risk-taking behavior, even terrorist practitioners put a certain premium on their own lives. However, as studies by Sprinzak and a highly influential piece by Pape (2003) observed, what may seem irrational on an individual level may be quite logical at the group level. That is, terrorist groups may benefit from the use of individual members to conduct suicide attacks, whereas the individual conducting the attack may not.

In addition to the psychological impact suicide terrorism causes, the tactic offers several advantages for groups willing to engage in extreme violence. First, in terms of material cost, suicide terrorism is cheap. The amount of explosive needed to rig a human bomb is small, and the monetary costs for acquiring the components are minimal to the group. Second, the operational complexity of planning a suicide attack is greatly reduced when there is no need for an exit strategy for the attacker. Third, the suicide attack is, on average, more deadly than any other form of terrorism. Pape's most controversial argument concerned the effect of suicide attacks. He argued that the reason suicide attacks increased in popularity is that terrorist groups observed that they were successful in gaining territorial concessions from democratic states. For example, Hezbollah successfully drove the U.S. Marines and French paratroopers out of Lebanon following two massive suicide bombings. The use of suicide terrorism by the LTTE coerced the Sri Lankan government into establishing an autonomous region for the ethnic Tamils in the early 1990s. Israel abandoned the Gaza Strip and West Bank in the mid-1990s as a consequence of being targeted by suicide bombers. Pape believed that democracies are susceptible to the effects of suicide bombing because, unlike autocracies, they are accountable to a public that recoils at such attacks. The free and open media that are characteristic of a democracy are also better able to publicize attacks than are their state-controlled counterparts in autocracies.

There are some shortcomings with Pape's work. Some of the targeted nations characterized as democracies in his

analysis did not possess convincing democratic credentials. There are also cases of suicide attacks directed at decidedly authoritarian regimes, such as the Saudi monarchy. Furthermore, because of the closed nature of the media in authoritarian states, many examples of suicide bombings may have gone unreported. Ultimately, though, Pape's argument can be extended to terrorism in general: It is popular because it sometimes works in achieving the goals of the groups that use it.

Conclusion

This chapter has illuminated several aspects of the form of political violence known as terrorism. The phenomenon differs from other forms of political violence. It has occurred within many different periods, regions, and cultures and has been executed under a variety of ideological auspices. The theoretical and empirical contributions of various social scientists have advanced the study of terrorism into an academic subfield with significant explanatory power.

References and Further Readings

- Associated Press. (2004). Analysis: Only U.S. can damage self majorly despite Bin Laden. *USA Today*. Retrieved April 22, 2010, from <http://www.usatoday.com/news/world/2004-12-18-bin-laden-x.htm?csp=34>
- Blomberg, S. B., Hess, G. D., & Orphanides, A. (2004). The macroeconomic consequences of terrorism. *Journal of Monetary Economics*, 51, 1007–1032.
- Bueno de Mesquita, E., & Dickson, E. S. (2007). The propaganda of the deed: Terrorism, counterterrorism, and mobilization. *American Journal of Political Science*, 51(2), 364–381.
- Crenshaw, M. (1981). The causes of terrorism. *Comparative Politics*, 13(4), 379–399.
- Enders, W., & Sandler, T. (1991). Causality between transnational terrorism and tourism: The case of Spain. *Studies in Conflict & Terrorism*, 14(1), 49–58.
- Enders, W., & Sandler, T. (2006). *The political economy of terrorism*. New York: Cambridge University Press.
- Enders, W., Sandler, T., & Parise, G. F. (1992). An econometric analysis of the impact of terrorism and tourism. *Kyklo*, 45, 531–554.
- Frontline. (2007). *Timeline: Struggling to find a strategy for success in Iraq*. Retrieved April 22, 2010, from <http://www.pbs.org/wgbh/pages/frontline/endgame/cron/>
- Hassan, N. (2001). An arsenal of believers. *The New Yorker*, 77(36), 36.
- Hoffman, B. (1998). *Inside terrorism*. New York: Columbia University Press.
- Jenkins, B. (1974). *International terrorism: A new kind of warfare*. The RAND Corporation. Retrieved April 22, 2010, from <http://www.rand.org/pubs/papers/2008/P5261.pdf>
- Krueger, A. B., & Maleckova, J. (2003). Education, poverty and terrorism: Is there a causal connection? *Journal of Economic Perspectives*, 17(4), 119–144.
- Lake, D. (2002). Rational extremism: Understanding terrorism in the 21st century. *Dialog IO*, 1, 263–296.
- Laqueur, W. (2001). *A history of terrorism*. New Brunswick, NJ: Transaction.
- Mason, T. D., & Krane, D. A. (1989). The political economy of death squads: Towards a theory of the impact of state sanctioned terror. *International Studies Quarterly*, 33, 175–198.
- Merari, A. (1993). Terrorism as a strategy of insurgency. *Terrorism & Political Violence*, 5(4), 213–251.
- National Commission on Terrorist Attacks Upon the United States. (2004). *Complete 9/11 Commission Report*. Retrieved April 22, 2010, from <http://govinfo.library.unt.edu/911/report/index.htm>
- National Public Radio. (2003). *16th Street Baptist Church bombing: 40 years later; Birmingham still struggles with a violent past*. Retrieved April 22, 2010, from <http://www.npr.org/templates/story/story.php?storyId=1431932>
- Pape, R. A. (2003). The strategic logic of suicide terrorism. *American Political Science Review*, 97(3), 343–361.
- Sprinzak, E. (2000, September/October). Rational fanatics. *Foreign Policy*, 120, 66–73.
- Tilly, C. (2004). Terror, terrorism, terrorists. *Sociological Theory*, 22(1), 5–13.
- U.S. Federal Emergency Management Agency. (2009). *General information about terrorism*. Retrieved April 22, 2010, from <http://www.fema.gov/hazard/terrorism/info.shtm>
- U.S. State Department. (2001). *Patterns of global terrorism*. Retrieved April 22, 2010, from <http://www.state.gov/s/ct/rls/crt/2000/2419.htm>
- White, J. R. (1998). *Terrorism: An introduction*. Belmont, CA: Wadsworth.

15

POLITICAL AND MILITARY COUPS

CHRISTOPHER D. MOORE

Bethel University

In the study of comparative politics and international relations, few phenomena are quite as mysterious as the *coup d'état*. In many countries, coups are a regular and even frequent source of regime change. Coups can be observed nearly every year all around the world, and there is a large and diverse body of coups, successful and unsuccessful, for scholars to study. Yet fundamental questions about coups persist. In a revolution, we can observe mass social movements of people seeking change. In legal regime changes, whether hereditary or democratic, we can observe a process that is frequently clear and usually public that guides the transition of power from one government to the next. This clarity is absent in most coups d'état. This is problematic for both scholars and political leaders. Scholars want to understand coups in order to grasp the overall picture of regime change in international politics. Political leaders need to understand coups as they are likely to encounter this sudden and unpredictable change in government, possibly even their own.

In an effort to explore this puzzling phenomenon, numerous scholars have attempted to explain the nature, causes, and consequences of coups. This chapter first defines and characterizes coups d'état. Next, it discusses the main factors thought to cause coups to emerge. Third, it examines the consequences of coups, both within the state and for other interested countries. A necessary component of this examination is a discussion of how states seek to

prevent coups from occurring. Last, this chapter suggests some directions for future research on coups d'état.

The Nature and Causes of Coups d'État

What is a coup? Like many terms in the study of international politics, the term *coup* has entered our popular lexicon. Coups have come to refer to any action that is characterized as sudden, decisive, dramatic, and usually highly successful. As we shall see, some of these characteristics, though certainly not all, typify coups d'état. The word *coup* is of French origin and means *stroke* or *blow*. Therefore, a coup d'état is literally a "blow of state" or, more idiomatically, a blow aimed at or against a state. However, this tells us little about the nature of coups and what distinguishes coups d'état from other forms of regime change, especially their close cousin, the revolution. The first task of this section, therefore, is to provide a definition of coups d'état.

Definition of Coups d'État

Broadly, a coup d'état is a form of regime change. When coups are successful, a new or at least partially new government supplants the prior power structure. Coups are similar in this way to all other forms of regime change, from democratic elections to mass revolutions. There are,

however, key differences that distinguish coups from other types of regime change. These key differences are discussed below.

Extralegal or Illegal Transfer of Power

Perhaps most obviously, a coup d'état is not a legal transfer of power. If a transfer of power occurs according to some legislative action or constitutional directive, a coup has not taken place. It is tempting to say that coups are irregular or abnormal transfers of power. However, coups happen so regularly and repeatedly in some countries that this is not a useful way of understanding coups d'état. Furthermore, notions of irregularity or abnormality create a status quo bias in our understanding of coups d'état. Coups may or may not lead to better governance for their people, but by treating them as irregular or abnormal transitions of power, we give undue benefit to the leadership previously in power. Therefore, it is useful to think of coups as regime changes occurring outside of any stated or previously agreed-on transfer of power.

The Noncritical Role of Violence

In comparison with other forms of regime change, coups d'état do not necessarily rely on violence or the threat of violence to affect the transition of power to a new government. This notion may initially seem counterintuitive, but once understood, it helps to clearly differentiate coups d'état from other forms of regime change. In a revolution, the government is expelled from power by direct violence or threat of violence. In a coup d'état, the government loses power because an elite faction is able to use its resources to acquire a preponderance of power in government without significant violence or coercion.

In casual discussions of the topic, the term *bloodless coup* has become popular. This term is illustrative of the general nature of coups, even if the literal application of the term is misleading. Coups are most successful when the least amount of physical violence or threat of violence must be marshaled. This is because coups leverage the already existing power of the cooperating elites to wrest control of the full political apparatus for themselves. Consider the prototypical example of a military coup. A group of generals grows dissatisfied with the way the civilian government is prosecuting a war. Rather than develop a revolutionary force to fight the government, the generals use their governmentally vested power to command their soldiers to surround the capital for the purposes of protecting it. Now, with their power well displayed, the generals can assert authority over the civilian leadership in the capital. If the coup is optimally successful, the civilian leadership will turn its power over to the generals. If the generals encounter resistance, publicly arresting or executing a few resisters demonstrates both power and restraint. Resistance can be crushed with a minimum number of casualties. This

is clearly a different political process from mass revolution, which often leads to substantial casualties (Leiden & Schmitt, 1968). Therefore, it is clear that coups d'état reflect a type of regime change in which the importance of violence is tangential at most.

The Inconsequential Nature of Popular Will

In the modern world, most regime change occurs as the result of some form of popular will. In the case of democratic elections, an electorate of citizens or subjects selects its new leadership. Popular will also matters in revolutions and civil wars. In civil wars, the people will tend to back a new government when it meets their ideological interests (because it is fighting for causes they support) or because they believe it is in their material interest to do so. That is, they believe that the new government is better able to provide for their security or that their support will mitigate punitive action from the new government. Either way, public support helps secure the transition of governing power after a civil war has concluded.

In the case of revolutions, popular will is even more directly involved. Popular will is a basic requirement for a popular uprising, and whether the uprising gains sufficient strength to depose the current government depends largely on the extent of the popular will supporting it. Social movements seek policy change, and when the favored policy change can best be accomplished through regime change, that social movement becomes revolutionary in character. Therefore, the popular will of the movement empowers the revolution. Without popular support, the revolution will fail.

This requirement is not at all necessary for coups d'état. Coups d'état regularly seem to occur without any consideration of popular will. One important scholar of coups, Edward Luttwak (1979), affirms this characteristic of coups in his seminal work, *Coup d'Etat, a Practical Handbook*. For a coup to be successful, Luttwak believes, the smallest number of functionaries needed to seize the reins of power is the preferred coalition. In fact, the coup is frequently a tool of those looking to contradict the popular will in the case of a democratic election. Whether Luttwak's notion is correct or not, it is clear that the coup d'état is fundamentally an elite phenomenon rather than a product of popular will. The elites who most often seem to be involved in coups are governmental actors, military actors, and economic leaders (Jackman, 1976). As discussed below, the coup's elite character serves as a key in both predicting the occurrence of coups d'état and designing state policies to prevent them.

It should be noted that just because coups originate from small groups of elite individuals, popular support is not perpetually irrelevant. A coup can seize power, but once it begins to govern, its continued success depends on popular support. Coups that fail to capture popular support typically fail (Luttwak, 1979). This may also help to explain why the

presence of a coup is one of the most powerful predictors of subsequent coups, a factor also discussed below (Bueno de Mesquita, Siverson, & Woller, 1992).

Affirmation of Social and Institutional Structures

Some authors have argued that coups d'état fall within a category of political phenomena known as *elite political instability* (Fosu, 2002). This term seems to suggest that coups are destabilizing to a social system. In effect, this is true. Coups, and especially repeated coups, do seem to negatively affect economic growth. However, it is more important to note that the intended effect of a coup is not to destabilize the social or political institutional structures of a society. This sets the coup d'état apart from other forms of social movements or revolutions (Gurr, 1970; Kousis & Tilly, 2005; Tilly, 1978; Tilly & Tarrow, 2007).

Consider the brief example above of the military coup deposing a civilian government. The generals have no interest of invalidating their own authorization to command their troops. Doing so would undermine their own power. The last thing the instigators of a coup d'état would like to do is fundamentally upset the current social order, because they are elites in it. Rather, their interest is in emphasizing social and political institutions that allow them to claim power. In this way, coups are by nature more conservative than other forms of regime change. This helps us understand why some coups come to power promising policy change whereas others come to power promising to prevent policy change. The coup is not necessarily progressive or reactionary but rather a lateral transfer of power within the existing power structure of a society.

Types of Coups d'État

The four criteria discussed above form the groundwork for a definition of the coup d'état:

A coup d'état is an extralegal transfer of power that affirms traditional social and political power structures and occurs without major contributions from violence or popular will.

However, within this definition, there is still a great deal of room for different types of coups. The differences tend to revolve around the nature of the participants in the coup and their proximate objectives. For example, many scholars distinguish coups that originate through civilian forces from those that originate within a state's military. The military's role in the coup d'état is studied more frequently than any other portion of the state apparatus (First, 1970; Jackman, 1976; Jackman, O'Kane, Johnson, McGowan, & Slater, 1986).

This distinction is further subdivided by Samuel Huntington, who describes three types of military coups (Huntington, 1968). Huntington (1968) describes a *guardian coup*, which seeks to replace ineffective governance with effective governance. This coup fits the military

generals' actions described above. In contrast, a *break through coup* is attempted by junior officers and seeks to undermine the military hierarchy, using civilian power to validate the transition of power. Finally, a *veto coup* is used by the military to prevent action taken by the civilian government against the interests of the civilian military. Other authors add additional types of coups, including the *self coup*, in which a government attempts to seize additional powers for itself. This conception of coup is problematic because if the legality of the power seizure is contested, the coup can merely be considered part of politics as usual.

The Causes of Coups d'État

Now that a definition of coups d'état has been established, we can turn to the pressing issue of causality. Like many political phenomena, the incidence of coups d'état appears to be multicausal. This section will briefly describe some of the causes of coups and conclude by evaluating their relative importance in the general phenomenon of coup origins. For conceptual clarity, the proposed causes of coups are subdivided into acute and chronic conditions. These categories are not unproblematic, but they reflect the notion that some proposed causes are proximate to the origin of the coups, while others are societal conditions that set the stage for a coup to emerge. These terms are used instead of direct and indirect causes of coups in an attempt to demonstrate that all the proposed variables are seen as probabilistic, rather than deterministic, causes of coups.

Acute Causes of Coups d'État

Many proposed causes of coups are acute, in that they occur proximate to the origins of the coup itself and might be seen as part of the process that directly leads to the coup itself. A number of these causes stem from factors within the military. Eric Nordlinger (1977) discusses the role of grievances within the military as a source of willingness to challenge the civilian government. These grievances may be particularly strong if they combine structural grievances with heightened personal grievances on the part of specific military leaders. Others consider the role of military popularity in the eyes of the public. Higher popularity may tempt elements within the military to seize power.

In addition to military causes, other domestic causes yield logical links to the emergence of coups. Luttwak (1979) points to domestic political crises as a window of opportunity for the emergence of a coup. These political crises could be constitutional, economic, or even electoral in nature.

Other acute causes of coups are generated from external sources. Michael Desch (1999) argues that military threats from abroad can create opportunities for coup instigators to obtain power, promising to ward off external threats. Bruce Bueno de Mesquita et al. (1992) find that military defeat and participation in war raise the likelihood of a coup. Finally, direct foreign intervention may increase the likelihood of a coup d'état. According to one study,

approximately 10% of all coups d'état involve the participation of a foreign power (David, 1987). There is no simple relationship to this involvement. In some cases, it appears that foreign involvement is the original cause of the coup attempt. This is typified in Operation Ajax, a U.S. effort that led to the deposition of Mohammed Mosaddeq in Iran in 1953. In other cases, coup instigators reach out to foreign powers in an attempt to gain additional power and leverage to apply domestically. The 1965 military coup that seized power in Algeria appears to have done this by reaching out to other regional powers in Africa (Quandt, 1969).

Chronic Causes of Coups d'État

There are numerous structural and institutional factors associated with the emergence of coups d'état. Aaron Belkin and Evan Schofer (2003) have developed a structural model of coup risk, which serves as a very useful overview. The first set of factors revolves around the material or structural factors of the state in question. First, what is the character of the economy of the state? If a state's export economy is particularly monolithic, this factor will hypothetically raise the likelihood of a coup. If a single industry or export is particularly powerful or influential, then controlling the management of that industry would be critical to capturing the power of the state. Certain types of exports lend themselves to this type of consolidation at the export level. Oil, valuable minerals, timber, and other materials that are of high value and require capital-intensive extraction are particularly vulnerable to this type of coup propensity.

A second structural condition is centralization of wealth (Jackman et al., 1986). This factor is correlated with the notion of a monolithic export economy described above. If wealth is highly centralized, coup risk is hypothetically elevated because the coalition of wealthy elites necessary to cause a coup is considerably smaller. An egalitarian distribution of wealth should have the effect of decreasing the likelihood of a coup d'état.

In addition to economic variables, scholars have considered historical factors that may influence coup propensity. Key among these are the lingering effects of colonialism. Many former colonies have well-developed and well-trained military forces. Often, the level of institutionalization and cohesion in considerably higher in the military than in the civilian government. This was the case in Pakistan when Pervez Musharraf was able to seize power in 1999. Musharraf was able to gain power not only because the military held violent coercive power but because the military was understood to be as legitimate a basis of power as the civilian government of Nawaz Sharif. Colonial powers such as France and Great Britain often emphasized the training of military forces over the establishment of indigenous civilian institutions. This choice may serve to make former colonies more prone to coups. In addition, colonial legacies may affect the full enfranchisement of the citizenry. Huntington (1968) points to this idea in proposing that coups are more likely in societies that limit political

engagement to a subset of their citizenry. Citizens outside the political process may seek a coup as a means of political participation. Note, however, this hypothesis somewhat contradicts the idea that coups are primarily elite driven.

Other historical factors affecting the likelihood of coups d'état are the legitimacy of the regime in power and the strength of civil society. Even if the military or other organs of the state are highly cohesive and institutionalized, if there is a correspondingly strong civil society, coups are much less likely to occur (Jackman, 1978). A strong civil society raises the minimum number of participants necessary to carry out a coup, and if the society is strong enough, the number of necessary coup participants will rise beyond the point of feasibility. Additionally, if the regime in power is seen as legitimate, Nordlinger (1977) argues, a coup will be less likely. Some have argued that this variable is close to tautological as those deposing a regime of course will consider it illegitimate.

A final historical factor that influences the likelihood of a coup is a previous coup. Numerous authors point out that coups often beget additional coups (Bueno de Mesquita et al., 1992). There are several logical ways to explain this empirical finding. First, the presence of a coup may be an indicator of societal conditions that promote the emergence of a coup. Since a coup by nature does not seek to alter societal or institutional power structures, the factors that led to the first coup persist and may allow for additional coups. A second explanation can be called the *political instability pathway*. Because coups by nature are extralegal transfers of power, a coup weakens the legitimacy of social and political institutions. In the same way, then, that a heart attack weakens the muscles of the heart, a coup weakens the institutions of power. This creates windows of opportunity for additional coups. Finally, some authors point out that new regimes are generally less stable than established, institutionalized regimes. Therefore, any time a regime changes, by coup or otherwise, it is more likely to be susceptible to instability, including additional coups (Sanderson, 2005).

What can we learn from these diverse hypothetical causes? Clearly, a portrait of a state ripe for coup d'état is beginning to emerge. Such a state is perhaps a former colony; has a monolithic export economy; exhibits weak governmental legitimacy; and has a weak civil society, a clear economic elite, and a powerful, institutionalized, and cohesive military. When these factors are coupled with military grievances, external threats, or domestic political crises, a coup is likely.

Internal and External Policy Implications of Coups

While political scientists have expended a great deal of effort to define and consider the causes of political and military coups, a growing area of study has been the consequences of coups d'état and the policies that states implement to deal

with the potential for coups both at home and abroad. This section thus deals with three questions. First, what are the consequences of coups d'état? Second, how do states attempt to protect themselves against the onset of a coup, and how successful are these efforts? Third, how do states react to coups in other countries, and what role does this reaction play in their foreign policy?

Consequences of Coups d'État

After a political or military coup d'état has taken place, several effects can be regularly observed. Before we can consider any others, we must turn to the success of the coup. Coups d'état regularly fail. In a study of 16 West African countries from 1955 to 2004, Patrick McGowan (2006) found that the military plotted a coup 169 times. Out of these plots, 87 led to actual coup attempts. Of these attempts, 43 failed and 44 succeeded. So the success rate of a coup depends on when measurement begins. Looking only at coup attempts, we would find that coups succeed about half the time. However, given our theoretical understanding of coups, we would expect the success rate to be quite high because coup instigators would initiate coups only when their chances of success were very high. According to this logic, a 50% success rate is low, assuming that the costs for the instigators are quite high. Alternatively, we could measure coups from the origins of a serious plot. By this measure, coups succeed only a quarter of the time. This seems more appropriate as many coup plotters will be dissuaded by the difficulty of successful execution. This observation raises an important question in the study of coups d'état: Why are coups initiated in the face of such high costs when they fail so frequently? A relevant direction for future research is discussed in the final section of this chapter.

When coups do succeed, we first can observe the steps that the coup instigators undertake to consolidate their power. These steps generally fall into two categories: *institutionalization* and *integration*. The institutionalization step reflects the attempts of the coup instigators to legitimize and normalize their new position as the leadership of the state. The integration step reflects their attempts to buttress their position in leadership by making additional bargains with power centers inside and outside the state in order to prevent their removal. Whereas revolutions and democratically elected governments enter with a mandate of popular will, coup instigators do not enjoy this benefit. As a result, the days immediately following a coup are the most critical for the new leadership. It is in this time that the new leaders are most susceptible to a countercoup or other action that will remove them from power.

Coup instigators begin to institutionalize their leadership through interaction with elites and the general populace. The new leadership often makes use of its current sources of power and attempts to spread that legitimacy throughout the society (Leiden & Schmitt, 1968). For example, when Musharraf seized power in Pakistan, his chief source of power was the Pakistani military he controlled. Because the

military of Pakistan is highly institutionalized, deeply professionalized, and politically active, Musharraf was able to use this position to convey a sense of legitimacy to both the Pakistani populace and other Pakistani elites. Often accompanying this portrayal of legitimacy is a set of promises or other forms of political framing to communicate a justifying narrative for the seizure of power. For example, the new leaders may claim that they seized power to end the corruption of the prior regime. They may also seize power to restore the state to greatness, to prevent calamity, to end oppression, or for other noble causes. Several prominent scholars refer to these promises as the coup's *pronouncements* (Luttwak, 1979; Nordlinger, 1977). Whether anyone deeply believes these tales may not be important. So long as the new leadership is able to prevent a countercoup, its pronouncements become part of the political narrative.

Coup instigators also work to integrate their power throughout the state they have seized. While the institutionalization step might be dismissed by some as mere political theater, the integration step is clearly crucial to the short-term success of any coup. When a coup succeeds, by definition a small number of individuals place themselves at the top of the governance structure. For their effort to be successful, they must link themselves fully to the traditional reins of power in the government and society. A helpful analogy is to think of a heart transplant. In the days and months following a heart transplant, recipients must typically take medication designed to prevent the body from rejecting the new organ. The new leadership must initiate a similar process after a coup. The leadership must reach out to other leaders, social institutions, industries, and other sources of power and convince them that they can trust the new government and may even find it advantageous. This behavior takes the form of reassurance, deal making, and incorporation. The ideal, for the coup instigators, would be to reassure these power sources in society that the new government would not dramatically alter the status quo. Often, this coincides with the coup's political rhetoric: The coup itself may be promising to return the society to a previously understood status quo. However, when powerful players in a country's domestic politics cannot be reassured, the new leaders may offer deals. Deals can take the form of increased political power or autonomy for cooperation with the new regime. This form of deal making can lead to the final option of the new leadership, *inclusion*. In order to bolster its own power after a coup, the new leadership may be forced to invite other power brokers from the society into the new government. Obviously, this is not advantageous for the coup instigators, and it may ultimately lead them to be more susceptible to a secondary coup by the very powers they are forced to include by necessity in their new government.

"Coup-Proofing" and Other Internal State Policies

When we consider the causes of coups d'état, we are immediately presented with a list of hypotheses for the

prevention of coups. However, most literature on the prevention of coups has focused on two topics: the limited number of structural conditions that generally protect states from coup-type instability and the broader set of conscious policies some states undertake to reduce or eliminate the likelihood of future coups in their country.

First, a substantial number of studies have emerged out of political economy to test what structural factors ward off the emergence of a coup. In seeking to understand coup risk, many of these studies grapple with Huntington's (1957, 1962) contention that modernity is the best defense against a coup, but that modernization is a great risk for political instability. If this is true, is it also true that less developed states, those likely to be susceptible to a coup, are unable to move into a low-risk environment without jeopardizing their own political stability? According to John Londregan and Keith Poole (1990), not exactly. Testing a large set of countries' economic and political behavior from 1950 to 1982, Londregan and Poole find several results that confirm the causes of coups mentioned above. Particularly potent among these causes are the existence of a prior coup and poverty within a country. However, Londregan and Poole also find that not only is a higher rate of societal wealth a powerful inhibitor of future coups, but economic growth is also a surprisingly effective coup preventative. In this respect, studies of development tend to disagree with Huntington; economic modernization does not necessarily lead to a coup as long as societal income is increasing (Jackman et al., 1986). This suggests that all types of governments, including authoritarian governments or those seizing power via a coup d'état, should have an interest in promoting economic development as a means of insulating themselves against a future coup.

In fact, states will go to great lengths to prevent coups. Not surprisingly, states that have suffered coups or whose governments came to power as the result of a coup d'état often exhibit the most conscious policies that attempt to shield the state against future coups. When the government of a state cannot count on economic development as a means to protect itself against a future coup, it may adopt several types of policies, which can be organized into categories of redundancy, surveillance, loyalty, and proper funding. These concepts are well summarized and illustrated in several recent works, notably that of James Quinlivan (1999).

Redundancy

One of the chief weapons of all states against a coup is redundancy. Because coups come to power by seizing control of a critical aspect of the state apparatus, one means of guarding against this is to create multiple versions of state institutions. Most often, this behavior is seen in the military of a state and the state's intelligence services. First, many states, and particularly states with reason to fear future coups, will create multiple militaries with parallel hierarchies. This is particularly true in authoritarian governments.

In Saddam Hussein's Iraq, the Revolutionary Guard served as a separate, parallel military hierarchy to the normal Iraqi army. The creation of the Revolutionary Guard helped to insulate Saddam's Baathist regime against potential coups from the regular military apparatus. At the same time, the regular military served as a check against coups emanating from the Revolutionary Guard.

States often also employ redundancy in their intelligence services. Because intelligence services by nature work with restricted information and clandestine policies, they are often breeding grounds for coups d'état as a form of regime change. A prime means of preventing this is the creation of multiple intelligence agencies, whose mission in part includes spying on each other. By positioning itself as the sole arbiter between warring intelligence services, the leadership of the state assures itself that it holds the most complete intelligence picture and that its intelligence services are too busy with turf battles to challenge its leadership.

Surveillance

A second powerful policy of coup-proofing is surveillance. While intelligence services regularly gather information on subversive groups within a state, surveillance speaks to a broader sense of scrutiny that inhibits the emergence of coups. Here, the analogy of Jeremy Bentham's *panopticon* is useful. According to Bentham, the panopticon was a prison with a sphere of cells surrounding a single watchtower, whose guard the prisoners could not see. Because prisoners could always potentially be observed, and because they could never know whether they were actually being observed, Bentham believed that their behavior would be scrupulous. If the powerful members of a society function in an environment where trust is low and there is an ongoing belief in their continued surveillance of each other, the opportunities for forming a successful coup coalition are very few. Some leaders, through public trials of corruption or treason, will demonstrate that the society or at least the government is vigilant against threats. As long as the perception within society is one of general vigilance and relatively low trust, many potential coup conspiracies can be dissuaded, even if they were unlikely to be detected in the first place.

Loyalty

A third and very powerful tool of coup-proofing is to rely on previously existing forms of loyalty. In many societies fearful of political instability, there are often deep ethnic or religious cleavages. Governmental leaders will capitalize on these social cleavages as a means of insulating themselves from threats of a coup d'état. By promoting members of their own religious sect or ethnic group into positions of power throughout the society, the governmental leaders are assured that challenges to their authority are reduced. A coup that displaces them would also displace the leadership of other power centers in the society. This means

that a coup has fewer places to generate from in the first place. Quinlivan (1999) points to this type of ethnic and religious networking in such supposedly coup-proof states as Saudi Arabia, Syria, and Saddam Hussein's Iraq.

Funding and Appeasement

As was stated above, economic development and income growth are powerful inhibitors of coups d'état. However, not every state continuously experiences such prosperity. When income is not growing in a society or economic development is not occurring, the leadership of the state may worry that the risk of a coup is increased. When this is the case, one of the most direct responses that a state can take is to ensure that those it supports financially do not challenge its authority. This means that the state must make certain that elements within the state apparatus are economically satisfied. If the state fears that a coup might originate from the military, the risk of such a coup is reduced if the military is well paid and well compensated (Byman, 2006). When the vast majority of individuals within an organization are individually satisfied, the pool of potential coup conspirators is very shallow. This observation also serves as a corollary to the idea described above that disastrous foreign conflicts could be a source of potential coups. When the military is insufficiently compensated or is losing a war, its dissatisfaction may raise the risk of a coup. Therefore, keeping critical organs of the state satisfied will help ward off the onset of a coup.

Foreign Policy Responses to External Coups

There are two general approaches to considering the foreign policy implications of coups. The first approach is to suggest that because coups are deeply idiosyncratic, foreign reactions to coups will be highly varied, with very little predictability from coup to coup. While it is certainly true that each coup is unique in its full range of participants, execution, and consequences, arguing that there are no systematic reactions to coups d'état is also misleading. Therefore, this section presents the broad but observable foreign policy reactions that coups d'état elicit from other states. These responses can generally be further subdivided into foreign policy responses toward the coup state and foreign policy responses because of the coup itself.

Responses Directed Toward the Coup State

When states react to a coup d'état within another state, two factors seem to predominantly affect the response. The first factor is domestic politics within the state. Democratic states usually respond negatively to coups that subvert the democratic will of another people. This has been aptly illustrated in the U.S. reaction to the recent military coup in Honduras. Much of the intransigence of the U.S. reaction can be explained by noting that it is not clear who is defending the democracy of Honduras. Is it Manuel

Zelaya, the democratically elected leader who was deposed by military forces? On the other hand, were the military forces seeking to preserve democracy from Zelaya, who appeared to be looking to extend his rule through constitutional manipulation? It seems that should the United States have been able to arrive at a decisive answer to this question, it might have adopted a position in favor of one side or the other. However, as of this writing, the United States continues to push for compromise between Zelaya and the military, suggesting that it sees democratic intentions on both sides of the contest.

Ideology also played a major role in external coup support or opposition during the cold war. Numerous examples of coups engendered by the United States and the Soviet Union in third-party countries suggest that the superpowers were not only willing to support military and civilian coups d'état to obtain more favorable state leadership but also to help create these coups. Particularly instructive on the U.S. side are the cases of Guatemala and Chile (Cullather, 2006; Linz & Stepan, 1978; Westad, 2007). In these cases, the United States went against democratic inclinations to support coups d'état that replaced governments that looked as if they might favor the Soviet Union.

Ethnic motivations may also play a role in responses to coups d'état. When the coup is executed by forces characterized by shared ethnicity, Donald Horowitz (1985, 2001) finds, we can expect that the coup's leaders will receive support not only from domestic ethnic compatriots but also from international kinsmen. Horowitz also describes a *seesaw coup*, whereby different ethnic groups struggle for control of a country through coups and countercoups, with each side receiving external as well as internal ethnic support.

Generally, when a state opposes a coup in another country, whether for ethnic or for ideological reasons, the typical response is to sever diplomatic ties with the new government and to refuse to recognize its legitimacy. If the new government can survive, its survival raises problems for the opposed state. At what point will the opposed state reengage with the government? Often, democratic countries tie their reengagement to some form of democratic activity. For example, if the new government promises or delivers democratic elections, this strategy serves as an avenue toward rapprochement. On the other hand, when a foreign government favors a coup d'état, it is often complicit in the institutionalization and integration project of the new government. It is in the foreign government's interest to legitimize the new, coup-installed government as quickly as possible.

Responses Stemming From the Coup Itself

Although most of the observable foreign policy of states in reaction to a coup d'état are directed toward the coup itself, one particular form of reaction is directed internally. This chapter has already discussed the extent to

which a previous coup raises the likelihood of a future coup. Richard Li and William Thompson (1975) find that states also react in ways that suggest that they worry that a coup in a proximate country might raise the risk of a coup in their own country. Li and Thompson find that in fact the presence of a coup in a nearby country does elevate the risk of coup in one's own country. They conclude that this process is a result of reinforcement, whereby coup instigators are emboldened by coups occurring in other, nearby countries. As a consequence, many states that witness nearby coups often are more likely to adopt coup-proofing behaviors of the types described above. It is also possible, although presently untested, that coup contagion is actually a product of self-fulfilling prophecy. When a state reacts to a nearby coup with coup-proofing policies, the inception of these policies themselves may actually increase the likelihood of a coup d'état.

Future Directions for the Study of Coups d'État

The systematic study of coups d'état has, unfortunately, given way in recent years to investigations of other types of regime change. Part of the space often given to the study of military and political coups has been overtaken by a renewed interest in terrorism and complex insurgencies. This change in scholarly pursuits may reflect generational preferences, or it may reflect the changing nature of international politics. The dust of the third great wave of democratization has settled, and there is no longer a cold war to impel some coups forward. However, even a casual accounting of recent coups suggests that coups d'état remain an important and relevant phenomenon of international politics. It is important that recent authors have also demonstrated that the risk of a coup can dramatically influence multinational foreign policy. Daniel Byman (2006) has argued that a coup in a country such as Saudi Arabia, Pakistan, or Egypt could dramatically inhibit U.S. efforts to fight al Qaeda in the Middle East. The study of coups remains important, and the importance will grow as scholars turn to answer the following questions.

First, is there a simple definition of coup success? There remains substantial disagreement within scholarship on coups d'état on the critical factors determining coup success. Some argue that the critical factor is the execution of the coup. If a coup succeeds in supplanting the previous government, then the coup is considered successful (Luttwak, 1979; Nordlinger, 1977). In contrast, coups are politically successful only if the coup instigators remain in power long enough to develop rudimentary institutionalization and integration and subsequently exercise power. This cannot really happen in the early days of a coup, but how long must a coup endure before it is considered successful?

Second, are coups discrete phenomena or contextual events? Many large quantitative studies of coups d'état treat them as discrete phenomena. In some ways, there is

much to recommend this choice. It suits the data requirements of statistical studies. Coups are clandestine and focused on a small group of instigators, so they appear to emerge from nowhere. However, this chapter has detailed a number of ways that coups affect and are affected by structural and historical conditions. The context of the coup matters, and this context is well treated by the idiographic historical studies of coups. Unfortunately, too many histories of coups treat them as idiosyncratic, with no knowledge to be gained by a comparative study of coups d'état. It seems that future studies of coups must learn from both these perspectives. Future studies need to take great account of the social and historical factors influencing coups while at the same time seeking a nomothetic, comparative understanding of coups as a social phenomenon.

Third, what is the normative place of coups? The scholarship on military and political coups d'état has rarely stopped to consider the ethical questions associated with a nondemocratic regime change. Future research should examine the normative implications of coups. While coups are attempted with a variety of intentions, do these intentions matter? If a military coup seeks to save the country or prevent broader pain to its people, do these motives make a coup acceptable? A second line of normative questioning relates to violence. No coup is truly bloodless, in the sense that someone must be coerced out of power. However, some modern coups seem to be nearly bloodless. Does this matter normatively? Should it change how foreign states ought to respond to the coup? A final consideration is the role of popular will. What if, as may have been the case in Honduras, a coup is reflective of popular public opinion? Does that factor justify the use of extralegal, nondemocratic means to seize power? Future research should consider these questions and integrate them into a comparative analysis of coups d'état.

A fourth and final set of questions is methodological. What theoretical approaches and methodological tools wait to be applied to the study of coups? One theoretical approach that has not yet been applied to the study of coups d'état is *political psychology*. Political psychology can add to our understanding of coups d'état by helping future scholars theorize when coups d'état are accepted, the role of perception and framing for the legitimacy of the coup instigators, and the role of risk propensity in initiating a coup. For example, could the extensive work on decision making under conditions of risk help future scholars understand why some coups are initiated and others are not (Farnham, 1994; McDermott, 1998)? A second theoretical approach is the voluminous and continually growing social movement literature. Although many classic studies of coups d'état make use of social movement literature, important advancements, both theoretical and methodological, can be applied to the study of coups (Davis, 2005; Tilly & Tarrow, 2007). In addition to new theoretical directions, there are opportunities for new methodological directions in the study of coups. Rather than relying on large sample, quantitative studies or historical case studies, the

phenomenon of coups could be studied through the use of *process tracing*. As opposed to correlation, which is the main finding of statistical studies, process tracing seeks to use nuanced historical data to test hypotheses within a single case. Combining this method with a quantitative study could yield more insightful results. A second fruitful methodological path is to employ computer simulations of regime change. Recent advances in computer simulation software allow for increasingly complex forms of behavior to be modeled and iterated in a game-theoretic context. This methodology would be a useful way to further test hypotheses in areas such as coup-proofing and coup contagion.

Conclusion

This chapter has sought to explain the state of scholarship on political and military coups d'état. Coups are an extralegal transfer of power that affirms traditional social and political power structures and occurs without major contributions from violence or popular will. Among acute causes of coups, foreign intervention and military defeat help explain the emergence of coups. Among chronic causes, the most powerful predictor is the presence of a prior coup, although economic and political factors change coup risk propensity. Certain policies tend to follow coups as the new regime seeks to consolidate its power. Other policies seek to prevent future coups through a process known as coup-proofing. There are unanswered questions, untapped theoretical traditions, and underused methodological traditions that will keep the study of coups rich and productive for years to come.

References and Further Readings

- Belkin, A., & Schofer, E. (2003). Toward a structural understanding of coup risk. *Journal of Conflict Resolution*, 47(5), 594–620.
- Bueno de Mesquita, B., Siverson, R. M., & Woller, G. (1992). War and the fate of regimes: A comparative analysis. *American Political Science Review*, 86, 638–646.
- Byman, D. L. (2006). Friends like these: Counterinsurgency and the war on terrorism. *International Security*, 31(2), 79–115.
- Cullather, N. (2006). *Secret history: The CIA's classified account of its operations in Guatemala, 1952–1954* (2nd ed.). Stanford, CA: Stanford University Press.
- David, S. R. (1987). *Third world coups d'état and international security*. Baltimore: Johns Hopkins University Press.
- Davis, G. F. (2005). *Social movements and organization theory*. New York: Cambridge University Press.
- Desch, M. C. (1999). *Civilian control of the military: The changing security environment*. Baltimore: Johns Hopkins University Press.
- Farnham, B. (Ed.). (1994). *Avoiding losses/taking risks: Prospect theory and international conflict*. Ann Arbor: University of Michigan Press.
- First, R. (1970). *Power in Africa* (1st American ed.). New York: Pantheon Books.
- Fosu, A. K. (2002). Political instability and economic growth: Implications of coup events in sub Saharan Africa. *American Journal of Economics & Sociology*, 61(1), 20.
- Gurr, T. R. (1970). *Why men rebel*. Princeton, NJ: Princeton University Press.
- Horowitz, D. L. (1985). *Ethnic groups in conflict*. Berkeley: University of California Press.
- Horowitz, D. L. (2001). *The deadly ethnic riot*. Berkeley: University of California Press.
- Huntington, S. P. (1957). *The soldier and the state: The theory and politics of civil military relations*. New York: Vintage Books.
- Huntington, S. P. (Ed.). (1962). *Changing patterns of military politics*. New York: Free Press of Glencoe.
- Huntington, S. P. (1968). *Political order in changing societies*. New Haven, CT: Yale University Press.
- Jackman, R. W. (1976). Politicians in uniform: Military governments and social change in the third world. *American Political Science Review*, 70(4), 1078–1097.
- Jackman, R. W. (1978). The predictability of coups d'état: A model with African data. *American Political Science Review*, 72(4), 1262–1275.
- Jackman, R. W., O'Kane, R. H. T., Johnson, T. H., McGowan, P., & Slater, R. O. (1986). Explaining African coups d'état. *American Political Science Review*, 80(1), 225–249.
- Kousis, M., & Tilly, C. (Eds.). (2005). *Economic and political contention in comparative perspective*. Boulder, CO: Paradigm.
- Leiden, C., & Schmitt, K. M. (1968). *The politics of violence: Revolution in the modern world*. Englewood Cliffs, NJ: Prentice Hall.
- Li, R. P. Y., & Thompson, W. R. (1975). The “coup contagion” hypothesis. *Journal of Conflict Resolution*, 19(1), 63–88.
- Linz, J. J., & Stepan, A. C. (1978). *The breakdown of democratic regimes*. Baltimore: Johns Hopkins University Press.
- Londregan, J. B., & Poole, K. T. (1990). Poverty, the coup trap, and the seizure of executive power. *World Politics*, 42(2), 151–183.
- Luttwak, E. (1979). *Coup d'état, a practical handbook*. Cambridge, MA: Harvard University Press.
- McDermott, R. (1998). *Risk taking in international politics: Prospect theory in American foreign policy*. Ann Arbor: University of Michigan Press.
- McGowan, P. (2006). Coups and conflict in West Africa, 1955–2004: Part II, empirical findings. *Peace Research Abstracts Journal*, 43(3).
- Nordlinger, E. A. (1977). *Soldiers in politics: Military coups and governments*. Englewood Cliffs, NJ: Prentice Hall.
- Quandt, W. B. (1969). *Revolution and political leadership: Algeria, 1954–1968*. Cambridge: MIT Press.
- Quinlivan, J. (1999). Coup proofing: Its practice and consequences in the Middle East. *International Security*, 24(2), 35.
- Sanderson, S. K. (2005). *Revolutions: A worldwide introduction to political and social change* (1st ed.). Boulder, CO: Paradigm.
- Tilly, C. (1978). *From mobilization to revolution*. Reading, MA: Addison Wesley.
- Tilly, C., & Tarrow, S. G. (2007). *Contentious politics*. Boulder, CO: Paradigm.
- Westad, O. A. (2007). *The global cold war: Third world interventions and the making of our times* (1st paperback ed.). Cambridge, UK: Cambridge University Press.

16

RESOURCE SCARCITY AND POLITICAL CONFLICT

WALTER THOMAS CASEY II

University of Texas at Dallas

“Destroy the thrones of the wicked. . . .”¹

The popular press is rife with economists, ecologists, and religious doomsayers seeking to explain, predict, and profit from the problem of resource scarcity and its twin, resource allocation. Closely following are an increasing number of scholarly forays in economics, anthropology, geography, and political science.

The modern scholarship’s groundwork may be said to properly begin with Malthus’s *Essay on the Principle of Population* (1798). Therein, he in part argues that given finite resources, and an infinitely expanding population, political, social, and spiritual turmoil is inevitable. The flurry of activity the publication of that essay created has been ever present, so that subsequent generations of popular writers (e.g., Charles Dickens and others) were able to tap into the broader theme. Thus by the late 1800s, social-ist theorists were able to exploit Malthusian ideas as a means of broadcasting the desirability (or, following Marx, the inevitability) of democratic equality.

Neo-Malthusian conservationists such as Harry Overstreet (1915), Thomas Carver (1915), and Richard Ely (1916) reexamined the previous centuries’ analyses and added their own Edwardian twist—the need to conserve resources for “civilized” peoples. Yet for all their

scholarly attempts, the conclusions were proscriptive rather than prescriptive and more than a bit race based.

It was not until the 1920s that deeper analyses were conducted and prescriptions for policymakers proffered. These Wilsonian internationalists stressed the need for necessary resources such as food, water, and fuel to be distributed or redistributed according to need rather than according to profits. More vocal academics, such as Powers in his 1928 *International Institutions: Formal Mechanisms for Dealing With Resource Conflicts*, was but one among many who saw the necessity of supranational actors as arbiters of dispute.

The economic crisis of the 1930s saw more and more attempts to move policy actors to embrace anticonflict measures of resource problems (Barnes & Field, 1933; Burns, 1934). Many saw within Hitler’s demand for *lebensraum* one country’s rather nakedly pragmatic attempt at a chauvinistic redress of conflict scarcity and were appalled by the possibility of such a tactic being embraced by similar actors elsewhere (Speier, 1939; Spengler, 1937). World War II was to make those fears realized.

During the cold war era, academic attempts to be prescriptive about solving political conflicts over resource

scarcity either veered into ideological territory by asserting the necessity of the “free world” to triumph over the forces of communism² or were besmirched with the label of communism,³ as Galtung’s (1965) work was so tagged. This retarded the progress made earlier, and it was not until the later 1960s and 1970s, as researchers became more interdisciplinary and reached out to the physical and life sciences, that scholarship began to move forward again. Richard Cooper’s (1973) foray still stands as a paragon of this type of research.

Cooper’s now classic work on the economic anthropology and settlement patterns of the Hmong in northern Thailand was a breakthrough. By establishing how the production of opium in highland family units and the fights over such production affected them, Cooper melded anthropology, environmental science, and sociology in a much admired and much copied recipe for the study of resource allocation and policy reactions to potential conflicts. The weakness, of course, is in inferring generalizations from the individual level to larger actors. But nonetheless, it was a grand experiment in its observation of actualities, rather than an ivory tower research exercise. With Cooper’s research, resource allocations become more personal and more pointed, and hit closer to home.

Once research put a human face on the various aspects of resource scarcity and conflict, attempts to generalize became more possible because numerous field studies became less anecdotal as they were gathered together and compiled and became more empirical as they became data. Not all data are created equal, and many of the individual biases of the researchers remain in some of that gathered materiel. But data-gathering techniques have improved and become more uniform, and coding has proceeded apace. Thus anthropology, sociology, physical science, life science, geography, economics, and statistics come together in modern political science analyses.

The current state of scholarship can be best understood as a multidisciplinary effort, with political science having wrested much control of the discourse from other disciplines, due in no small part to the ability to meld, merge, and synthesize, as well as to create, novel empirical explorations. The following discussion attempts to thematically display the current state of scholarship.

The Current State of Research

Much of the literature rests on pure physical resource scarcity and allocation. This body of work assumes, predicated on historical events, that resources are finite or near finite and that one entity’s possessing an amount of a resource necessarily denies access, production, or use of that resource to another entity.⁴

Ross’s (2004) review of cross-national econometric and qualitative studies of scarcity and conflict concludes that

collectively, most prior work can be grouped according to its conclusory assumptions:

1. Oil increases the likelihood of conflict, particularly separatist conflict.
2. “Lootable” commodities such as gemstones and drugs do *not* make conflict more likely to begin but do prolong existing conflicts.
3. There is no apparent link between agricultural commodities and civil war.
4. The association between primary commodities, which include both oil and agricultural goods, and the onset of conflict is not robust.

Ross suggests that the inconsistencies among studies may be caused by differences in the ways researchers code civil wars and cope with missing data. His conclusion is all the more interesting given the problems inherent in qualitative studies that have been well documented along these lines. His conclusion thus seemingly implies the need for more rigor in empirical work and better cross-study coder reliability and pasigraphization of definitions and terms.

Finding little support for almost all empirical assessments, however, is the work of Urdal (2005). Noting that neo-Malthusians’ assertion that population pressure on natural renewable resources makes societies more prone to low-intensity civil war enjoys little support, he goes on to demonstrate that “resource-optimists’” assumption that agricultural land scarcity caused by increasing population density drives economic development, in turn driving peace, is unfounded as well. He tests both in a time-series cross section from 1950 to 2000, with results for either being tepid, at best, or counterfactual.

Countries experiencing high rates of population growth, high rates of urbanization, or large refugee populations do not face greater risks of internal armed conflict than other countries do. There is some indication that scarcity of potential cropland may have a pacifying effect.

However, where land scarcity combines with high rates of population growth, the risk of armed conflict increases somewhat. . . . Overall, the robustness of the empirical support for both paradigms is low. A strong emphasis on security as a macro rationale for reducing global population growth thus seems unwarranted. (Urdal, 2005, p. 417, italics added)

Becsi and Lahiri (2007) examine scarcity and conflict via a simple *trade theoretic model* (two regions in conflict, war equilibrium determined endogenously). Their finding is that an abundance of uncontested natural resources discourages conflict (a more even ratio) and an abundance of contested natural resources encourages conflict. They also note that when entities (regions) possess power to affect the terms of trade, and use that power, the effect of ownership on conflict “may be strengthened or weakened depending on factor intensities

of production and the relative strength of income and substitution effects” (p. 17).

Le Billon analyzes the vulnerability of resource-dependent countries, concludes that long-term stability in resource-exporting regions is dependent on their levels of development, and calls for a broad reform agenda prioritizing the basic needs and security of local populations. Similarly, Bannon and Collier (2003) find that the entities (in this case, nations) most likely to be affected by conflict are those whose economies depend mostly on natural resources, a phenomenon they term *resource dependence conflict*.

Bogalea, Taebb, and Endoc (2006) use multinomial discrete choice models to identify determinants of household choice among alternative land property right regimes and whether those rights help mitigate negative consequences of scarcity-induced land-related conflicts. They assert that two factors, dependency ratio (i.e., that proportion of a population composed of dependents, or people who are too young or too old to work. The dependency ratio is equal to the number of individuals younger than 15 or older than 64, divided by the number of individuals aged 15 to 64, expressed as a percentage) and level of education, produce the most predictive power, whether or not the household prefers a common property resource regime. Conversely, the variables *number of household members*, *livestock holding of the household*, and *area of cultivated land lost due to enclosure* were found to be more relevant in determining a preference for resettlement. Such a study assists the researcher in determining model framework and methodological choice, especially given the personal nature such discrete choices entail.

Lujala, Gleditsch, and Gilmore (2005), in examining “lootable resources” (here, diamonds), offer a competing model along similarly strong methodological lines. While territory, oil, and water are most often mentioned as the type of contestable resources likely to lead to conflict, diamonds have emerged in recent literature as a prominent factor. The authors find a strong bivariate relationship between diamonds (particularly secondary diamonds) and the onset of conflict, but adding diamond dummies to standard models produces less robust conclusions. The secondary–primary distinction is important because production of secondary diamonds increases the risk of ethnic conflict, but not other types, whereas primary diamonds make ethnic conflict less likely. Most intriguing is that the impact of diamonds has been substantially stronger in the post–cold war era, suggesting a tantalizing ideology–resource link.

Klare (2002) broadens the single-item resource paradigm to a more widely encompassing one. As the complexities of rapidly increasing demand of globalizing industrialization continue, the concentration of resources in unstable states and the competing claims to ownership of resources by neighboring states predict a greater likelihood of conflict. Examples include the potential for conflict

over oil in the Persian Gulf and in the Caspian and South China Seas; over water in the Nile Basin and other multinational river systems; and over timber, gems, and minerals from Borneo to Sierra Leone. Klare’s analyses of likely conflicts are informed by detailed research into projected usage rates, population growth, and other relevant trends that show such to affect the likelihood of conflict: a pattern repeated throughout the world.

The discourse on diamonds points up not only the break from primary resources to produced or manufactured resources but the ecological effect such production has as well. Joseph Stiglitz’s works (e.g., his 2002 *Globalization and Its Discontents*) have been a lodestone for those examining production and manufacture as a source of conflict. In contrast with earlier studies focusing on fuel, mineral, and agricultural resources and the efficiency by which they are allocated, the concern Stiglitz examines is similar to that of the climate-change scholars, in that he also includes the environmental consequences of resource extraction and use.

A New Variable: Climate Change

This new debate, that of human activity—particularly industrialization and its attendant problems—causing global climate change, is notable because an increasing number of players in this discursus have argued that climate change will lead to resource competition, mass migration, and, ultimately, an increase in armed conflict around the world. Salehyan (2008) takes issue with such “determinism.” He posits instead that the effect of climate change on armed conflict interacts with a number of political and social variables. Ignoring interaction effects results in spurious correlations and problematic predictions about when and where conflict is likely. He offers the assiduous scholar a research program predicated on more rigorous and sophisticated methodologies than the comparative case study.⁵ *Biocapacity* (a measure of how biologically productive land is) and *ecological reserve* (a measure of the amount of land available for production) are also predictive of peace, but the community of scholars engaged in this research arena cautions such results are less robust. Given the extractive elements necessary for possessing such an *ecological footprint*—which, for all intents and purposes, comprises the old idea of the first world—their critics assert that there exist conflating, interactive, and indirect effects not fully explained.

These interaction effects, which Homer-Dixon (2009) has noted since 1991,⁶ are ameliorated, he says, by the ability of the extractors, producers, and consumers to use “human ingenuity” positively. That said, this enfant terrible of this arena as well is willing to lay fault at the feet of crass corporate and elitist carpetbagging and the social structures of repression, as well as comment on

the role urbanization plays in exacerbating scarcity, as he has done since 1999. He does not, in the end, insist that ingenuity can or will more than partially offset scarcity-related problems, which only reductions in scarcity can do fully.

This finding is similar in vein to what Gausset, Whyte, and Birch-Thomsen (2005) attempt. They deliver an alternative perspective to conflict analysis by building on theories of political ecology that have developed from Marxist geography and cultural anthropology. Such an approach sets aside the neo-Malthusian accounts anchored in population determinism and linear causality. They instead ask for a neo-dialectical approach by focusing on the processes by which natural resources are manipulated by vested interests for assuming power, some arguments of which have been explored previously by Peluso and Watts (2001).

A weakness of the dialectical approach is in not acknowledging that “one of the important contributions of Malthusian and neo-Malthusian models is that they make individual agency central” (Peluso & Watts, 2001, p. 17). Glossing over this aspect of natural resource consumption and carrying capacity are problematic. A more appropriate model, Peluso and Watts (2001) argue, bridges the dialectic with historicity of study. To that end, they develop the concept of *political scarcity*, but this itself tends to neglect the physical constraints of resource availability.

The importance of political models as part of the study of scarcity is one that Nobel laureates such as Amartya Sen and Wangaari Maathai have commented on. But the danger of a political scarcity model is potentially absolving individual consumption patterns and ecological behavior of human societies. Moreover, although the political model approach is commendable, failing to account for material aspects suggests an incomplete model. Bailis (2006) notes the incompleteness of contemporary models, and Reuveny (2008) asserts that if explanation is part of empirical social science, so then is prediction. That is, the effects of climate change on migration can be forecast by exploring the effects of earlier environmental conditions on migration in recent decades. He argues that adaptation, mitigation, and emigration are the three possible human responses, the choice dependent on the extent of problems and mitigation capabilities. People living in lesser developed countries may be more likely to leave affected areas, which may cause conflict in receiving areas.

Stringer et al. (2009) agree but, as with most work in this area, use a regional study to support that theory. While not fully generalizable, their postdictive examination of climate-sensitive development strategies serves to make their foray an observational natural experiment of sorts—itsself very useful. Hendrix and Glaser (2007), examining a wider area (sub-Saharan Africa), contribute further by examining multiple time dimensions, both long-term trends in climate and short-term climatic triggers on civil

conflict onset. They find, similar to the first world operationalization noted earlier, that climate suitable for Eurasian agriculture is associated with a decreased likelihood of conflict, which is concomitant with distribution ratios of freshwater resources per capita being positively associated with the likelihood of conflict. Using simulated data up to 2099, they also predict climate conflict, finding, surprisingly, that there are few statistically significant, positive trends, suggesting the lesser probability of climate change dramatically affecting conflict. That in turn infers a Bayesian, or instantaneous learning and updating process, is at work intraculturally, or even cross-culturally.

Such is contra Martin, Blowers, and Boersema (2006). They begin by noting that conflict is often intratribal, -ethnic, or -religious. But large data sets disconfirm association between a country's social diversity and the likelihood of warfare; rather the converse relationship is more often found: ethnic and religious diversity tends to foster peace, not war. This departs from Crawford and Lipschutz's (1998) *Myth of "Ethnic Conflict."* Martin et al. assume two largely independent dimensions to the study of environmental conflict. The first investigates what has been dubbed the *resource curse*. Resource abundance (especially oil) is viewed as a factor in motivating secessionist movements, whereas a wider range of high-value resources (oil, gold, drugs, coltan, tin, diamonds, timbers) provides a means of financing rebel armies and thus sustaining and escalating already existing violence (Bannon & Collier, 2003). In contrast to concerns about abundance, resource scarcity may be viewed as a possible cause of violence. Scarcity in this context refers mainly to renewable resources such as water, fuelwood, and soil and arises from reduced supply (depletion or degradation), increased demand, increasing inequality of distribution, or a combination of these.

Brown, Hammill, and McLeman (2007) review the linkages between climate change and security in Africa and analyze the role of climate change adaptation policies in future conflict prevention. Ameliorating or adapting to such projections necessitates an accounting of the other variables interacting: existing social, political, and economic tensions.

In terms of indirect effects, Obioha (2008) asks what happens when *climate change* (herein, depressed rainfall) → *scarcity* (herein, crop shortages) → *conflict*? Focusing, as much of the literature does, on sub-Saharan Africa—Nigeria in particular—Obioha investigates the chain of interactions between climate change, population drift and pressure, and conflict over land resources. Thus he is able to use not only the indirect effects model but an admixture of indirect and interaction effects—not only a more sophisticated method but a more nuanced and inclusive one.

This is what Reuveny (2008) does as well, and across time and space. His foray into post-Katrina Louisiana

and Mississippi, 1950s Bangladesh, and the dust bowl of the United States during the 1930s offers a model that demonstrates pre- and postpolicy changes due to climate disasters and demonstrates that environmental change can trigger large outmigration, leading to violent conflict in areas receiving migrants. He asserts that policies seeking to minimize migration induced by climate change and violent conflict in receiving areas require an engineered economic slowdown in the developed countries and population stabilization and economic growth in the developing countries, financed by the developed countries. Either is a prescription for uneasy politics.

Within the realm of climate-change and political interaction models lie the works of those focusing on renewable, yet restricted, resources. Examples abound in examinations of water scarcity and conflict. The water conflict in northern Thailand was precipitated by a water shortage that ruined an orchard, causing the lowlanders to become angry toward the highlanders, as El Niño lessened precipitation during the rainy season of 1997. Ekkawatpanit, Kazama, Sawamoto, and Ranjan (2009) show that the conflict itself was less about a widespread water shortage (affecting most parties equally) than about the ability of wealthy farmers to purchase water while poorer ones became insolvent, which demonstrates that water scarcity on a discursive level does not reflect the real scarcity.

This is different from Pearce (2007), who makes a compelling case that a worldwide fresh and potable water shortage is *the* most fearful looming environmental crisis. His expository delivery of statistical evidence is indeed doomlike. What this work supplies the researcher is data, accompanied by a grim narrative. Unfortunately, the empirical connections are lost, and the effort comes across as less than rigorous.

Water, argues Shiva (2008), is intrinsically different from other resources and products and cannot be treated simply as a commodity: Without water, people and the environment cannot survive. To subject water to commercial restrictions and to control its availability to people and communities are normatively unacceptable. Contrary to others who claim that water scarcity will lead to conflicts in the future, Shiva provides evidence that water wars are already with us and are happening all over the world. She is convinced that conflicts will become increasingly violent as freshwater resources dwindle. Unfortunately for many readers, there is a tendency to dwell on the ethics and morals, and not fully develop the empirical arguments.

Humphreys, Sachs, and Stiglitz (2007) lead a post-modern indagation focusing on a particular extracted and processed resource, petroleum. They not only make a careful assessment of the effect extraction and processing have on climate but also discuss economic hindrances and helps deriving from oil funds. Their work is

the best representative of a host of efforts examining how the fight over petrodollars has exacerbated scarcity and conflict.⁷

Where Next?

The reassessments of this line of inquiry—specifically, methodological and empirical reassessments—have thus added new vigor to the discourse and sharpened the debate outside academe. An excellent take on how measurement matters comes recently from O’Lear and Diehl (2007):

In armed conflicts, particularly those involving natural resources and other environmental factors, the *issue of scale* remains overlooked and underanalysed. Although previous work has considered conflict at different analytical resolutions, scale itself is rarely addressed directly or as an important characteristic of a conflict. . . . Literature on natural resource related conflict has tended to overlook issues of scale and create a self imposed constraint on our understanding of conflict by determining a priori, and often indirectly, the scope or frame of a conflict. (p. 179)

In addition, authors are beginning to reevaluate Western notions of scarcity and its effects—for instance, Ember, Ember, Korotayev, and de Munck’s (2007) per-lustrative examination of fat and thin, noting that resource scarcity and valuing fatness in women are negatively associated when there is little or no food storage and *unrelated* when there is moderate or high storage. Such work helps to shape the field by reevaluating the marks and measures, assumptions and paradigms of the effect scarcity may have.

What Resource Problem?

Not all work in the field accepts that scarcity necessarily leads to conflict. Such an assumption has come under increasing scrutiny, leading to more, and frankly better, empirical assessments. Not all are supportive of the earlier findings—note Theisen (2008): “The theory relating civil violence to the degradation of natural resources receives only limited support” (p. 802). Theisen’s most recent work finds little support linking resource scarcity and civil conflict, but it replicates earlier findings on the importance of poverty, instability, and dependence of fuel exports, seemingly vindicating the free-market critics. Of note as well is Urdal’s (2005) assumption:

In the environmental security literature, great rural resource scarcity, causing rural to urban migration, is seen as an important source of violent conflict. . . . Urban disorder is primarily associated with a lack of consistent political institutions, economic shocks, and ongoing civil conflict. (p. 418)

Within the world of scholarship and popular political punditry, the counterarguments tend to follow eight lines, with some overlap:

1. There is no such thing as resource scarcity (Wright & Czelusta, 2004).
2. Resource scarcity may exist, but property rights ameliorate it (Mehta, 2007).
3. Resource scarcity may exist, but it is not the source of violent conflict (Dwyer & Oh, 1987).
4. Resource scarcity exists only in nondemocratic and/or corrupt states (Kalyuzhnova & Nygaard, 2008).
5. Resource scarcity may exist, but conflict management mitigates it (Himes, 2008).
6. Resource scarcity cannot exist where sufficient technological advances exist (Gowdy & Julia, 2007).
7. Resource scarcity may exist, but free markets prevent it (Horwitz, 2008).
8. Resource scarcity exists, but the data support no causality (Salehyan, 2008).

Outside academe, the counterarguments tend to be more strident:

1. There is no such thing as resource scarcity, and the theory that there is, is Marxist propaganda (Mills, 2008).
2. Resource scarcity doesn't exist, and the sources of violent conflict are Marxist redistributive policies (Lott, 2007).
3. Resource scarcity exists only in nondemocratic states because free markets prevent it from appearing in democratic ones (Corsi & Smith, 2005).
4. Resource scarcity may exist, but climate change and environmental degradation aren't the causes; corruption at the local level is (Singer & Avery, 2008).

The non-resource-scarcity literature is thin in most parts, but some of the academic critiques are worthy of further examination. But most of the non-resource-scarcity literature finds itself running headlong into the same wall repeatedly—the data and evidence in support of the various resource scarcity theories far outweigh those that do not. In the words of Krautkraemer (2005), “The general conclusion . . . is that technological progress has ameliorated the scarcity of natural resource commodities; but resource amenities have become more scarce, and it is unlikely that technology alone can remedy that” (p. 2).

Conclusion

Richard Matthew's summary (2008) of the state of the discipline is a clear, concise assessment of much of the scholarship, and the essays by Dobkowski and Wallimann (2002) can certainly serve as an ideal introduction to the subject. A number of works cited in this all-too-brief examination of the literature would serve as well as subject-specific overviews. But a unified treatment of the problems of resource scarcity and political conflict issuing from it is yet to be seen.

Part of this absence is due to the problems of definition and delineation, part to measurement, and part to empirical analyses. Le Billon (2007) makes a strong case, following the comparativists' admonishment, that scale matters—not only area, but amounts and impacts. Otherwise, we are comparing apples not to apples, or oranges, but to bricks. It thus stands that much, then, needs to be done in unifying the disparate treatments, universalizing the language, and standardizing the definitions and measures. Perhaps then can scholarship proceed and policy prescriptions begin.

Notes

1. Luke 1:52ff: *deposuit potentes de sede et exaltavit humiles*; that is, “[god] demolishes the thrones of the unjust and exalts the lowly.”

2. See Craig Duncan (1962), *Economic Geography*, 38(2), 113–121, or Lynn White (1962), *Medieval technology and social change*, Oxford University Press; also Tibor Scitovsky (1954), Two concepts of external economies, *Journal of Political Economy*, 17(1), 143–151.

3. See especially Raymond Mack and Richard Snyder (1957), The analysis of social conflict: Toward an overview and synthesis, *Journal of Conflict Resolution*, 1, 212–248, for a paradigmatic example.

4. A way to consider this approach, and the subapproaches over particular physical resources, is thus

$$X_{r1t1} : Y_{r1t1}, \quad (1)$$

where X and Y are entities (individuals, groups, nations), $r1$ is a resource, and t is time. Since the resource $r1$ is finite or near finite, it is necessarily contestable. As the ratio of possession, use, or production of $r1$ becomes more uneven or disparate, conflict is possible (P). Thus,

$$X_{\Delta \uparrow r1t1} \longrightarrow Y_{\Delta \downarrow r1t1} \rightarrow P[\text{conflict}]_{t+n}. \quad (2)$$

Amelioration of conflict is, crudely, possible when $r1$ is more evenly allocated, or when other resources ($r2, r3, \dots rk$) become available that act as substitute or trade items for $r1$.

$$X_{\Delta \uparrow r1t1} \longrightarrow Y_{\Delta \uparrow r2t1} \rightarrow -P[\text{conflict}]_{t+n} \quad (3)$$

Much of the work on scarcity and conflict thereby proceeds from these assumptions.

5. Compare this to the slightly earlier work by Binningsbø, de Soysa, and Gleditsch (2007). They test a general argument about the effects of resource scarcity by examining the most widely used measure of environmental sustainability: the ecological footprint. Contrary to the neo Malthusian argument, they find that entities with a heavier footprint possess greater propensities for peace.

6. Thomas Homer Dixon's (2009) prolific output on this subject is itself worthy of exploration.

7. See, to list a very few, Matthias Basedau and Jann Lay (2009), Resource curse or rentier peace? The ambiguous effects

of oil wealth and oil dependence on violent conflict, *Journal of Peace Research*, 46(3), 757–776; Hanne Fjelde (2009), Buying peace? oil, corruption and civil war, 1984–99, *Journal of Peace Research*, 47(2), 199–218; or Syed Murshed and Mohammed Tadjoeeddin (2009), Revisiting the greed and grievance explanations for violent Internal conflict, *Journal of International Development*, 21(1), 87–111.

References and Further Readings

- Alao, A. (2007). *Natural resources and conflict in Africa: The tragedy of endowment*. Rochester, NY: Rochester Press.
- Bailis, R. (2006). Climate change mitigation and sustainable development through carbon sequestration: Experiences in Latin America. *Energy for Sustainable Development*, 10(4), 74–87.
- Bannon, I., & Collier, P. (2003). *Natural resources and violent conflict: Options and actions*. Washington, DC: World Bank.
- Barnes, J., & Field, F. (1933). *Behind the Far Eastern conflict*. New York: American Council, Institute of Pacific Relations.
- Beesi, Z., & Lahiri, S. (2007). Conflict in the presence of arms trade: Can foreign aid reduce conflict? In S. Lahiri (Ed.), *Theory and practice of foreign aid* (pp. 3–17). Amsterdam: Elsevier.
- Binningsbø, H., de Soysa, I., & Gleditsch, N. (2007). Green giant or straw man? Environmental pressure and civil conflict, 1961–99. *Population and Environment*, 28(6), 337–353.
- Bogalea, A., Taebb, M., & Endoc, M. (2006). Land ownership and conflicts over the use of resources: Implication for household vulnerability in eastern Ethiopia. *Ecological Economics*, 58(1), 134–145.
- Brown, O., Hammill, A., & McLeman, R. (2007). Climate change as the new security threat: Implications for Africa. *International Affairs*, 83(6), 1141–1154.
- Burns, A. (1934). The first phase of the National Industrial Recovery Act, 1933. *Political Science Quarterly*, 49(2), 161–194.
- Carver, T. (1915) *Essays in social justice*. Cambridge, MA: Harvard University Press.
- Cooper, R. G. (1973). *Resource scarcity and the Hmong response: Patterns of settlement and economy in transition*. Singapore: Singapore University Press.
- Corsi, J. R., & Smith, C. R. (2005). *Black gold stranglehold: The myth of scarcity and the politics of oil*. Nashville, TN: WND Books.
- Crawford, B., & Lipschutz, D. (1998). *The myth of "ethnic conflict": Politics, economics and "cultural" violence*. Berkeley: University of California.
- Dalby, S. (2007). Geopolitical knowledge: Scale, method and the "Willie Sutton syndrome." *Geopolitics*, 12(1), 183–191.
- Dobkowski, M. N., & Wallimann, I. (2009). *On the edge of scarcity: Environment, resources, population, sustainability, and conflict* (2nd ed.). Syracuse, NY: Syracuse University Press.
- Dwyer, F., & Oh, S. (1987). Output sector munificence effects on the internal political economy of marketing channels. *Journal of Marketing Research*, 24(4), 347–358.
- Ekkawatpanit, C., Kazama, S., Sawamoto, M., & Ranjan, P. (2009). Assessment of water conflict in Mae Chaem River Basin, Northern Thailand. *Water International*, 34(2), 242–263.
- Ely, R. (1916). Russian land reform. *American Economic Review*, 6(1), 61–68.
- Ember, C., Ember, M., Korotayev, A., & de Munck, V. (2007). Valuing thinness or fatness in women: Reevaluating the effect of resource scarcity. *Evolution and Human Behavior*, 126(3), 257–270.
- Fjelde, H. (2009). Buying peace? Oil wealth, corruption and civil war, 1985–99. *Journal of Peace Research*, 46(2), 199–218.
- Galtung, J. (1965) Institutionalized conflict resolution: A theoretical paradigm. *Journal of Peace Research*, 2(4), 348–397.
- Gamage, G. B., & Boyle, C. (2008). Sustainable development: A review of progress, stagnation and potential. *International Journal of Sustainable Development*, 11(1), 45–60.
- Gausset, Q., Whyte, M., & Birch Thomsen, T. (Eds.). (2005). *Beyond territory and scarcity: Exploring conflicts over natural resource management*. Uppsala, Sweden: Nordic Africa Institute.
- Gowdy, J., & Julia, R. (2007). Technology and petroleum exhaustion: Evidence from two mega oilfields. *Energy*, 32, 1448–1458.
- Gurr, T. (1970). *Why men rebel*. Princeton, NJ: Princeton University Press.
- Hendrix, C. S., & Glaser, S. M. (2007). Trends and triggers: Climate, climate change and civil conflict in sub Saharan Africa. *Political Geography*, 26(2), 695–715.
- Hewstone, M., & Greenland, K. (2002). Intergroup conflict. *International Journal of Psychology*, 35(20), 136–144.
- Himes, J. (2008). *Conflict and conflict management*. Athens: University of Georgia Press.
- Homer Dixon, T. F. (2009). *Carbon shift: How the twin crises of oil depletion and climate change will define the future*. Toronto: Random House Canada.
- Horwitz, S. (2008). Economists and scarcity. *The Freeman*, 58(5), 14–16.
- Humphreys, M., Sachs, J. D., & Stiglitz, J. E. (2007). *Escaping the resource curse*. New York: Columbia University Press.
- Kalyuzhnova, Y., & Nygaard, C. (2008). State governance evolution in resource rich transition economies: An application to Russia and Kazakhstan. *Energy Policy*, 36(4), 1829–1842.
- Kaplan, R. (1994). The coming anarchy. *The Atlantic Monthly*, 273(2), 44–76.
- Klare, M. (2002). *Resource wars: The new landscape of global conflict*. New York: Holt Paperbacks.
- Krautkraemer, J. A. (2005). *Economics of natural resource scarcity: The state of the debate*. Resources for the Future. Discussion Paper 05 14. Retrieved April 27, 2010, from http://www.rff.org/Documents/RFF_DP_05_14.pdf
- Le Billon, P. (2007). Scales, chains and commodities: Mapping out resource wars. *Geopolitics*, 12(1), 200–205.
- Lott, J. (2007). *Freedomnomics: Why the free market works and other half baked theories don't*. Washington, DC: Regnery.
- Lujala, P., Gleditsch, N. P., & Gilmore, E. (2005). A diamond curse? Civil war and a lootable resource. *Journal of Conflict Resolution*, 49(4), 538–562.
- Martin, A., Blowers, A., & Boersema, J. (2006). Is environmental scarcity a cause of civil wars? *Journal of Integrative Environmental Sciences*, 3(1), 1–4e.
- Matthew, R. A. (2008). *Resource scarcity: Responding to the security challenge*. New York: International Peace Institute. Retrieved April 27, 2010, from <http://www.ipinst.org/media/pdf/publications/rscar0408.pdf>
- Mehta, L. (2007). Whose scarcity? Whose property? The case of water in western India. *Land Use Policy*, 24(4), 654–663.
- Mills, R. (2008). *The myth of the oil crisis: Overcoming the challenges of depletion, geopolitics, and global warming*. New York: Praeger.

- Obioha, E. E. (2008). Climate change, population drift and violent conflict over land resources in northeastern Nigeria. *Journal of Human Ecology*, 23(4), 311–324.
- O’Lear, S., & Diehl, P. F. (2007). Not drawn to scale: Research on resource and environmental conflict. *Geopolitics*, 12(1), 166–182.
- Overstreet, H. (1915). Philosophy and the new justice. *International Journal of Ethics*, 25(3), 277–291.
- Pearce, F. (2007). *When the rivers run dry: Water – the defining crisis of the twenty first century*. Boston: Beacon Press.
- Peluso, N., & Watts, M. (2001). *Violent environments*. Ithaca, NY: Cornell University Press.
- Percival, V., & Homer Dixon, T. (1998). The case of Rwanda. In T. Homer Dixon & J. Blitt (Eds.), *Ecoviolence: Links among environment, population and security* (pp. 201–222). Lanham, MD: Rowman & Littlefield.
- Raleigh, C., & Urdal, H. (2007). Climate change, environmental degradation and armed conflict. *Political Geography*, 26(6), 674–694.
- Reuveny, R. (2008). Ecomigration and violent conflict: Case studies and public policy implications. *Human Ecology*, 36(1), 1–13.
- Ross, M. L. (2004). What do we know about natural resources and civil war? *Journal of Peace Research*, 41(3), 337–356.
- Salehyan, I. (2008). From climate change to conflict? No consensus yet. *Journal of Peace Research*, 45(3), 315–326.
- Schwartz, D., Deligiannis, T., & Homer Dixon, T. (2000, Summer). The environment and violent conflict: A response to Gleditsch’s critique and suggestions for future research. *Environmental Change & Security Project Report*, 6, 77–93. Retrieved April 28, 2010, from [http://www.homerdixon.com/download/Response to Gleditsch.pdf](http://www.homerdixon.com/download/Response%20to%20Gleditsch.pdf)
- Shiva, V. (2008). *Soil not oil*. Boston: South End Press.
- Singer, S. F., & Avery, D. T. (2008). *Unstoppable global warming: Every 1,500 years*. Lanham, MD: Rowman & Littlefield.
- Speier, H. (1939). Class structure and “Total War.” *American Sociological Review*, 4(3), 370–380.
- Spengler, O. (1937). *The decline of the West*. New York: Alfred A. Knopf.
- Stiglitz, J. (2002). *Globalization and its discontents*. New York: W. W. Norton.
- Stringer, L. C., Dyer, J. C., Reed, M. S., Dougill, A. J., Twyman, C., & Mkwambisi, D. (2009). Adaptations to climate change, drought and desertification: Local insights to enhance policy in southern Africa. *Environmental Science & Policy*, 12(7), 748–765.
- Theisen, O. M. (2008). Blood and soil? Resource scarcity and internal armed conflict revisited. *Journal of Peace Research*, 45(6), 801–818.
- Toman, M. A., Ayres, R. U., & Simpson, R. D. (2005). *Scarcity and growth revisited: Natural resources and the environment in the new millennium*. Washington, DC: Resources for the Future.
- Urdal, H. (2005). People vs. Malthus: Population pressure, environmental degradation, and armed conflict revisited. *Journal of Peace Research*, 42(4), 417–434.
- Wright, G., & Czelusta, J. (2004). Why economies slow: The myth of the resource curse. *Challenge*, 47(2), 6–38.

17

ETHNIC CONFLICT

TINA KEMPIN REUTER

Christopher Newport University

Ethnic conflict is one of the major threats to international peace and security. The conflicts in the Balkans, Rwanda, Chechnya, Iraq, Israel/Palestine, Indonesia, Sri Lanka, India, and Darfur are only among the best-known and deadliest examples. The destabilization of provinces, states, and in some cases even whole regions are common consequences of ethnic violence. Ethnic conflicts are often accompanied by gross human rights violations such as genocide and crimes against humanity, economic decline, state failure, environmental problems, and refugee flows. Violent ethnic conflict leads to tremendous human suffering.

Despite the fact that the number of conflicts has declined over the past decades, ethnic turmoil remains one of the main sources of warfare and instability in major regions of the world. Between 1945 and 1990, nearly 100 ethnic groups were involved in violent conflicts. During the 1990s, about three quarters of conflicts were disputes between politically organized ethnic groups and governments. More than one third of the world's states were directly affected by serious internal warfare at some time during the 1990s, and of these states, nearly two thirds experienced armed conflicts for 7 years or longer during the decade. In 2006, all 32 ongoing conflicts were internal, 5 of which were internationalized; most of them were caused by ethnic issues (Harbom & Wallensteen, 2007).

Ethnic Identity, Ethnicity, and Ethnic Groups

The terms *ethnic* and *ethnicity* have their roots in the Greek word *ethnos*, which describes a community of common descent. In ethnic conflict research, the terms *ethnic group*, *communal group*, *ethnic community*, *peoples*, and *minority* are mostly used interchangeably. Two elements provide the basis to identify ethnic groups: first, the accentuation of cultural traits, and second, the sense that these traits distinguish the group from the members of the society who do not share the differentiating characteristic. These *ethnic criteria*, which provide the origins of communal identity, may include shared historical experiences and memories, myths of common descent, a common culture and ethnicity (including race), and a link with a historic territory or a homeland (which the group may or may not currently inhabit). Elements of common culture include language, religion, laws, customs, institutions, dress, music, crafts, architecture, and even food. Ethnic communities show signs of solidarity and self-awareness, which are often expressed by the name the group gives itself (Smith, 1986). The definitions of the terms *ethnic* and *ethnicity* in ethnic conflict research thus go beyond the general usage in North America, where *ethnicity* commonly refers to race (skin color and other physical markers) only.

Ethnic identity is formed by both tangible and intangible characteristics. Tangible characteristics such as shared culture or race are important because they contribute to the group's feeling of identity, solidarity, and uniqueness. As a result, the group considers perceived and real threats to its tangible characteristics as risks to its identity. If the group takes steps to confront the threat, ethnicity becomes politicized, and the group becomes a political actor by virtue of its shared identity. On the other side, ethnicity is just as much based on intangible factors, namely, on what people believe, or are made to believe, to create a sense of solidarity among members of a particular ethnic group and to exclude those who are not (Smith, 1991) members.

Although communal identity provides the foundation for the definition of ethnic groups, disagreement exists over how ethnic identity forms and how it changes over time. A first school of thought, known as the *primordialist approach*, explains ethnicity as a fixed characteristic of individuals and communities (Geertz, 1973; Isaacs, 1975; Smith, 1986). According to primordialists, ethnicity is rooted in inherited biological traits and/or a long history of practicing cultural differences. Ethnic identity is seen as unique in intensity and durability and as an existential factor defining individual self-identification and communal distinctiveness. Mobilization of ethnic identity and ethnic nationalism is a powerful tool to engage the group in a political struggle. Ethnic divisions and ethnic conflict are considered inherent to multiethnic societies and a common phenomenon.

The primordialist focus on fixed identities, however, fails to recognize variation in ethnic group formation, ranging from relatively short-term associations to long-standing, strong, and cohesive groups with biological and historical roots. To account for these differences, a second, so-called *instrumentalist*, approach developed, which understands ethnicity as a tool used by individuals and groups to unify, organize, and mobilize populations to achieve larger goals (Brass, 1985; Glazer & Moynihan, 1975; Noel, 1968). These goals are mostly of a political nature and include, among others, demands for self-governance, autonomy, access to resources and power, respect for the group's identity and culture, and minority rights. In this view, ethnicity has little or no independent standing outside the political process and is in its character comparable to other political affiliations such as ideological beliefs or party membership. According to instrumentalists, ethnicity is a result of personal choice and mostly independent from the situational context or the presence of cultural and biological traits. Ethnic conflict arises if ethnic groups compete for the same goal, notably power, access to resources, or territory. Elite interest plays an important role in mobilizing ethnic groups to engage in ethnic conflicts. Ethnic conflict is thus similar to other political interest conflicts.

Critics of instrumentalism argue that ethnicity, in contrast to political affiliations, cannot be decided on by individuals at will but is embedded within and controlled by the

society as a whole. Advocates of *social constructivism* point to the social nature of ethnic identity and argue that ethnicity can only be understood in a relational framework (Anderson, 1991; Brubaker, 1995; Dominguez, 1989; Laitin, 1986). In their view, ethnicity is neither fixed nor completely open. Ethnic identity is created by social interactions between individuals and groups and remains therefore beyond a person's choice, but it is subject to change if the social conditions change. Individuals and groups cannot escape the fact that ethnic differences exist, but they determine themselves what they make of these differences (Wolff, 2006). Ethnic conflict depends thus to a great extent on the opportunities provided for the group to reach their goals. Violent conflict is caused mainly by social and political systems that lead to inequality and grievances and do not offer options for the peaceful expression of differences (e.g., discriminatory regimes). Changes in social interactions, such as increased tensions or violent conflict, influence the socially constructed nature of ethnicity. Social constructivists explain the tremendous atrocities committed during ethnic conflicts, such as genocide, mass rape, ethnic cleansing, and so forth, by the fact that by virtue of their ethnicity, everyone is part of the struggle (Chipman, 1993).

A fourth view ascribes to ethnicity deep cultural and psychological roots, which make ethnic identity extremely persistent (Ross, 2001; Volkan, 1997). *Psychocultural interpretations* stress the importance of shared, deeply rooted worldviews that shape group members' relationships with others, their actions and motives. These worldviews influence members' perception of origin, the intensity of their identity, and the significance of political action. Ethnic identity cannot be changed, only made more tolerant and open-minded. Ethnic conflict engages central elements of each group's identity and invokes fears and suspicion about real and potential opponents. Ethnic conflict is thus not simply a political event but a drama that challenges the very existence of the group by contesting its identity. This explains why ethnic conflicts are very difficult to resolve.

In reality, some ethnic groups have identities with deep historical roots whereas others do not, and some groups have static identities, whereas others have dynamic identities. The concrete expression of ethnicity and its propensity to lead to violence and warfare depend on the context. Ethnic identities are adaptable to and activated by unexpected threats and new opportunities. Ethnicity cannot be politicized unless an underlying core of memories, experience, or meaning moves people to collective action. As a result, ethnic identity usually "can be located on a spectrum between primordial historical continuities and (instrumental) opportunistic adaptations" (Esman, 1994, p. 14).

Several factors contribute to the salience and intensity of ethnic identities. Indisputably, the strongest factor is war and violence. First, the history of common efforts, stories of sacrifices for a common goal, and memories of human suffering create strong connections among the members of affected ethnic groups. Similarly, if a group

experiences economic, political, and cultural discrimination, group cohesion tends to increase. Second, a group's ethnic identity is stronger if mass literacy is achieved. Literacy allows elements of identity to be stored in writing, which means that historical and cultural narratives can reach a mass audience and stay the same over time. Even if an ethnic identity lies dormant for some time, it can be revived. Finally, the identities of nonimmigrant groups tend to be more pronounced than the identities of immigrant ethnic groups. While immigrants often assimilate, nonimmigrant minorities generally adhere to their traditions, especially if they are easily distinguished from the rest of the society by tangible traits such as physical markers (Gurr, 1993).

Not all ethnic groups are politically active or engage in ethnic conflict. According to the Minorities at Risk Project (www.cidcm.umd.edu/mar/), which tracks 283 mobilized ethnic groups, at least 17.4% of the world's population identifies with politically active ethnic groups. Depending on the political structure of the state (democracy vs. authoritarian regimes) and the size and situation of the ethnic minority (large vs. small portion of the society, regionally concentrated vs. dispersed), ethnic groups will have different claims and will use different means to voice their demands. The Minorities at Risk Project distinguishes six different group types: ethnonationalists, indigenous peoples, ethnoclasses, communal contenders, religious sects, and national minorities. *Ethnonationalists* are large, regionally concentrated ethnic groups with a history of autonomy or separatist struggles. Examples include the Quebecois in Canada, the Kurds in Iraq, and the Tibetans in China.

Indigenous peoples are original inhabitants of a colonized territory. These groups typically have traditional social, economic, and cultural customs that set them apart from the rest of the society (e.g., Native Americans, the Maasai in Africa, and the Aborigines in Australia). Even though indigenous peoples are often sharply distinct from the dominant group (they usually are set apart, not only by physical markers, but also by language, religion, traditions, etc.), they tend to be badly organized, have weak connections among group members, and, consequently, are usually unable to voice their claims (mostly to land and access to resources) in a successful manner. As a result, indigenous peoples are among the most marginalized ethnic groups in the world.

Ethnoclasses are racially or culturally distinct groups of people who are usually descendants from slaves or immigrants. African Americans in the United States or Muslim minorities in France are good examples of ethnoclasses. In many cases, these groups perform distinctive economic activities, mostly at the bottom of the economic hierarchy. Ethnoclasses generally strive for equal treatment, economic opportunities, and political participation. Mobilization of these groups varies widely. Ethnoclasses have successfully pursued their interests in many Western democracies whereas they remain relatively unorganized in most other places.

Communal contenders are culturally distinct groups that hold or seek a share in state power. Some of them can also be classified as ethnonationalists opting for separatism and seeking independence (e.g., the people of southern Sudan). The Minorities at Risk Project distinguishes between *dominant*, *advantaged*, and *disadvantaged* communal contenders. Dominant groups hold both political and economic power over other groups in their societies (e.g., the Sunni in Iraq under Saddam Hussein, Whites in South Africa during the apartheid regime, and the Tutsi in Burundi). Advantaged groups enjoy political benefits but are not in control of governing power (e.g., the Punjabis in Pakistan). Disadvantaged communal contenders are the most common; they often face political or economic discrimination or both (e.g., the Chinese in Malaysia and the Tajiks in Afghanistan). Changes to group relations involving communal contenders are particularly likely if power structures change. Intergroup shifts of relative political influence and economic prosperity can provoke violent actions, which tend to be particularly long lasting and disastrous, as illustrated by the conflicts between north and south Sudan (1956–2005) or different groups in Lebanon (1975–1990). Power-sharing models that take differences and external changes into account are the only way to deal with these issues. However, as history shows, these power-sharing arrangements are often very difficult to achieve.

Religious sects are ethnic groups that differ from the rest of the society, mostly by their religious beliefs and related cultural practices. Religious minorities tend to have high group cohesion because religion is a highly salient trait. In addition, religious groups usually already possess an organizational structure, which makes mobilization of the groups particularly easy and likely. Most groups in this category are Muslims and include both Islamic people in non-Muslim societies (e.g., Algerians in France, Arab citizens of Israel, or Turks in Germany) and different sects within a Muslim society (e.g., Sunni and Shi'a in Iraq). Non-Islamic groups include, among others, the Catholics in Northern Ireland, Jews in Argentina, the Copts in Egypt, and the Baha'i in Iran. For these politicized religious minorities, their faith is what sets them apart, but their goals are political in nature (e.g., participation in the government, nondiscrimination, or the recognition of the minority).

Finally, *national minorities* are groups with kinfolk in a neighboring state but who are a minority in the state in which they reside. Most of these groups have a history of political autonomy, which they strive to reinstate. Examples include Greeks in Albania, Russians in the Baltic, Hungarians in parts of Serbia, and Arabs in Iran.

The Origin and Nature of Ethnic Conflict

Conflict describes a situation in which two or more actors pursue incompatible goals. It is not necessarily violent, but the use of *tension*, *dispute*, or *unease* is more common in a nonviolent context. A violent internal conflict is generally

called a *civil war* or *armed conflict* if casualties and destruction are substantial, the conflict had a certain duration, the protagonists are organized, and military operations are used to achieve political goals (Brown, 2001b).

Ethnic conflict is a form of conflict in which the goals of at least one party are defined in ethnic terms, and the conflict, its causes, and potential remedies are perceived along ethnic lines (Horowitz, 1985). The conflict is usually not about ethnic differences themselves but over political, economic, social, cultural, or territorial matters. The conflicts in Northern Ireland or Israel/Palestine, for example, are not religious conflicts, but political conflicts, because the goals at stake are political, not religious in nature.

If the political goal of ethnic mobilization is self-determination, the movement is called *nationalism*. A *nation* in this context is a politicized ethnic group with the desire for self-government, ranging from participation in public affairs to local segmental autonomy to territorial claims, including independence (Van Evera, 1994). The use of the word *nation* is problematic. On the one side, *nation* can mean the state as a whole (the way the term is used in *international* or *United Nations*). If *nation* refers to people in this context, it can be understood as the aggregate, permanent population of the state, based on citizenship. On the other side, *nation* is also widely used to refer to a politicized ethnic group, in which case the link among people is based on ethnicity rather than citizenship.

Ethnic disputes are common in every multicultural society. Intergroup problems arise in periods of substantial political, economic, and social change and lead to uncertainty, emerging opportunities for action, and particularistic interests. Grievances and polarizing leadership lead to mobilization, ranging from political action (conventional politics, strikes, demonstrations, and other nonviolent means) to violent acts such as terrorism, armed uprisings, and guerrilla and civil wars (Horowitz, 2001).

Causes of Ethnic Conflict

Michael Brown (2001a, 2001b) distinguishes between *underlying* and *proximate* causes for ethnic conflict. Underlying causes include structural factors, political factors, economic and social factors, and cultural and perceptual factors. Proximate causes embrace four levels of conflict triggers: internal, mass-level factors (bad domestic problems); external, mass-level factors (bad neighborhoods); external, elite-level factors (bad neighbors); and internal, elite-level factors (bad leaders). Both underlying and proximate causes have to be present for ethnic conflict to evolve.

Underlying Causes

Structural Factors. *Weak states* or *failed states* are often a starting point for ethnic conflict. Most of these states are artificial products (e.g., former colonies) and lack

political legitimacy, ethnically sensible borders, and effective political and legal institutions. Violent conflicts are likely if changes in the economic situation of a state (e.g., cuts in foreign aid, corruption, administrative incompetence, and the inability to promote economic stability) are associated with the deterioration of the political situation in the country and the mobilization of ethnic groups. Group rivalry can lead to military mobilization, which leads to general armament of all ethnic groups within the state. This causes a security dilemma; by taking steps to defend themselves, ethnic groups often threaten the security of others (Posen, 1993). The ethnic security dilemma involves aspects of physical security (threats to the existence of the group), political security (oppressive regimes, exclusion from political participation), economic and social security (no equal opportunities for economic and social advancement of the group), cultural security (forced assimilation), and environmental security (destruction of a minority's land and resources; Wolff, 2006). Violent conflicts and internal security dilemmas lead to massive human rights violations, refugee flows, and spillover effects with the potential to destabilize whole regions.

Ethnic geography, namely, the geographic distribution and territorial concentration of ethnic groups in pluralistic states, is a second factor that contributes to the likelihood of violent ethnic conflict. Ethnic conflict is particularly common in states with territorially concentrated ethnic groups located near a border or with ethnic kin in an adjacent state (Fearon & Laitin, 2003). These groups show high levels of organization and increased group cohesion and are able to use shared homelands as a territorial base for their political struggle.

Political Factors. Ethnic conflict is particularly likely in states in which ethnic groups are inadequately represented in the government, the courts, the police, the military, political parties, and other public and political institutions. Authoritarian one-party regimes with discriminatory legislation and lack of opportunities for ethnic groups to participate in state decision-making processes are particularly prone to ethnic conflict. Liberal democracies that focus on the ideals of inclusion, political debate, and the attempt to reach consensus among all participants in the political process facilitate nonviolent ethnonational action and are thus less likely to experience rebellion or uprisings (Gurr & Harff, 2003). A second cause of conflict is exclusionary national ideologies. Nationalism and, in an increased form, citizenship based on ethnic distinctions are especially dangerous because such ideologies tend to flourish in situations of political uncertainty and economic collapse. Other forms of exclusionary national ideologies include religious fundamentalism and supremacist, fascist expressions. Third, the occurrence of violent ethnic conflict depends on stable domestic intergroup relations. Violent conflict is especially likely if the claims are incompatible, groups

are strong and organized, action is possible, success is achievable, and the fear of suppression and discrimination is tangible (Brown, 2001b). Tactics employed by leaders and elites during political turmoil are crucial: *Scapegoating*, *hate speech*, and *instrumentalization of the mass media* are means that have the potential to aggravate ethnic tensions.

Economic and Social Factors. Economic slowdowns, stagnation, deterioration, and collapse are sources of destabilization of the state and can lead to increased tensions and competition among ethnic groups. Competition for limited natural resources is one of the major factors leading to ethnic conflict. In addition, discriminatory economic systems with unequal economic opportunities, access to land and resources, and vast differences in standards of living generate resentment and contribute to tensions and destabilization. Fast economic transitions (e.g., from centrally planned to market economies) and development can aggravate instability by creating favorable conditions for domestic migration, urbanization, and other societal changes. These changes also raise hopes for economic and political gains that can provoke frustration if these expectations are not met.

Cultural or Perceptual Factors. Cultural factors such as problematic group histories, stereotypical perceptions, and grievances over cultural discrimination, including restricted educational opportunities, legal and political limitations on the use of the minority language, and constraints on religious and cultural practices, are common causes of ethnic conflict. In addition, a weakening of traditional forms of dispute settlement (such as a *council of elders*) changes the environment for conflict resolution of ethnic disputes (Brown, 2001a).

Proximate Causes

Proximate causes can be categorized according to (a) whether they are triggered by elite-level or mass-level factors and (b) whether they are triggered by internal or external developments. Brown (2001a, 2001b) identifies four main types of proximate causes of internal conflict:

1. Bad domestic problems (internal, mass level factor)
2. Bad neighborhoods (external, mass level factor)
3. Bad leaders (internal, elite level factor)
4. Bad neighbors (external, elite level factor)

First, internal mass-level factors create bad domestic problems such as rapid economic development, modernization, patterns of political or economic discrimination, and internal migration (urbanization). Refugees or fighters from neighboring countries who cross the border often bring violence and turmoil with them.

Second, radicalized politics can lead to contagion, diffusion, and spillover effects and create “bad neighborhoods”

(external mass-level causes). For instance, the Hutu refugee camps in Zaire became prime recruitment zones for rebel forces.

Third, internal elite-level aspects include power struggles by leaders of different groups, ideological contests over the way a country should be organized, and criminal assaults. Leaders have the ability to “play the ethnic card,” which can lead to increased tensions between ethnic groups. Milosevic’s policies in the former Yugoslavia are a good example. By using the national media, Milosevic fueled nationalist movements and hate toward non-Serbian groups, which led to ethnic cleansing and gross human rights violations committed during the wars in the 1990s.

And finally, external, elite-level factors are the results of decisions by governments to trigger conflicts in weak neighboring states for political, economic, security, or ideological reasons; an example is Russian involvement in Georgia (Abkhazia and South Ossetia). In addition, ethnic minorities in some cases decide to wage a violent struggle in the hope of political gains and international support. Ethnic groups assume the willingness of the international community to react and to provide a political forum to support negotiation, arbitration, and the settlement of disputes. The assumption of intervention by the international community can, in the worst case, cause the very tragedies international engagement in ethnic conflict tries to prevent. This happened, for example, in Kosovo in the late 1990s. The Kosovar Albanian rebel forces were convinced that if they could provoke the Serbs to attack ethnic Albanians, the international community would intervene on their behalf and thus facilitate their goal of independence. The plan seemed to work out: The rebels began shooting large numbers of Serbian police and civilians in 1997, the Serbs responded by bloody counterinsurgency in 1998, and the North Atlantic Treaty Organization bombed the Serbs in 1999, occupying the province and thereby establishing Kosovo’s de facto independence. However, both the Serb counterinsurgency and the Albanian attacks on Serbs after Serbia’s defeat caused the death and displacement of thousands of people on both sides, thereby leading to a tragedy that could have been prevented. These deaths were a direct consequence of the promise of humanitarian intervention (Kuperman, 2004).

Conflict Dynamics

Once ethnic conflict breaks out, it is difficult to stop. Massive human rights violations and physical attacks on civilians such as rape, torture, mass killings, ethnic cleansing, and genocide lead to tremendous human suffering. Systematic discrimination and exclusion from national and local political decision making, the appropriation of ethnic minorities’ traditional homelands, and policies that marginalize ethnic minorities are common practices accompanying ethnic conflict.

Even if fought at a low level of intensity, protracted ethnic conflicts have a great impact on the affected society. The lack of functioning or legitimate political institutions, weak economic performance, nonexistent or polarized structure of civil society, and antagonized elites lead to polarization and separation, eroding cross-cutting cleavages and leaving societies deeply divided and prone to further ethnic strife. In addition, ethnic conflicts have very direct effects far beyond their epicenters. These involve refugee flows, internal displacement, regional instability, economic failures, environmental disasters, diffusion and spillover effects, and conditions for organized crime and terrorism. Ethnic conflicts spread in two ways. Diffusion occurs when an ethnic conflict in one state stimulates conflict in another state with similar conditions. Successful movements provide images and moral incentives resulting in the motivation and mobilization of other ethnic movements in similar economic and political conditions. *Escalation* or *contagion effects* occur when a conflict in one country spreads across borders into neighboring countries in which an ethnic minority has its kinfolk. This usually involves the engagement of new foreign fighters who are employed by local elites. Ethnic conflicts may start out as intrastate disputes, but become regional or international crises when foreign powers get involved.

Neighboring states, regional powers, and international powers are often overwhelmed and unable to deal with international consequences of ethnic conflicts. However, in many cases, these external actors are not passive victims of ethnic crises but actively pursue their own agendas and interests. Foreign sympathizers and diasporas can contribute substantially to a group's cohesion and mobilization by providing financial, military, political, and moral support. External actors in some cases play important roles in inflaming conflicts or prolonging violent struggles. Opportunistic interventions to gain military, economic, or political benefits take advantage of conflict-affected states and contribute to the conflict. At the same time, international involvement can be crucial in preventing and settling ethnic conflict. The international community plays a role in negotiating, organizing, and supervising ceasefires and peace agreements; investigating past human rights violations; implementing the provisions of peace settlements; conducting peace operations including humanitarian, military, and economic assistance; imposing arms embargos and economic sanctions; and providing mechanisms of confidence and capacity building and of solving future disputes with peaceful means. Neighboring states and the international community can thus be victims of the troubles in the region or active contributors—sometimes deliberately, in other cases unintentionally—by providing military, economic, or political support of ethnic groups or engaging in negotiation and peace implementation. Regional instability is as much a source of ethnic conflict as it is a consequence.

Ethnic Groups and Ethnic Conflict Worldwide

Ethnic Groups

Given the aforementioned vague definition of ethnic groups, no one really knows how many ethnic groups exist in the world. Estimations range from a few hundred to a few thousand. The reasons for these discrepancies are manifold. Each state has different methods of determining group affiliation. While one state labels a group White, another distinguishes among different heritages. In addition, the fact that self-identification with an ethnic minority often comes with disadvantages leads to unreliable censuses. In other cases, ethnic minorities tend to overestimate their numbers to get benefits from the government or to put themselves in a stronger position (see, e.g., the entry for Albania in the U.S. Central Intelligence Agency's *World Factbook* (U.S. CIA, n.d.), which notes that as of 1989, estimates of the Greek population ranged from 1% in official Albanian statistics to 12% in statistics from a Greek organization). A further complication is that one ethnic group can have many different names. The group might have a name for itself, the state might have second one, ethnic kinfolk in a neighboring state might label themselves in a third way, and scholars might use a fourth name to refer to parts or the entire group. Finally, numbers fluctuate due to migration and other factors, such as fertility and mortality rates. According to the U.S. Census Bureau, for example, the growth of the U.S. Hispanic population from 9.6 million in 1970 to 102.6 million (projected) in 2050 will lead to major changes in the composition of the U.S. population, with the current majority (White) losing its majority status.

Around 80% of states are multiethnic societies, meaning that no ethnic group dominates the society. The remaining 20% are either states that are truly ethnically homogeneous (e.g., Japan and Korea) or states with overwhelming majorities (such as China, France, and Germany, which are home to many different ethnic groups). China, for example, has 57 official ethnic groups, but 91.5% of the people are Han (U.S. CIA, n.d.). In contrast, ethnically heterogeneous states comprise two or more ethnic groups, none of which is completely dominant. These groups can be regionally concentrated, as for example in Canada, Switzerland, or Belgium, or dispersed, as in the United States.

Ethnic Conflict

Ethnic conflict has been the world's most common source of warfare, insecurity, and loss of life. According to the Minorities at Risk database, 121 ethnic conflicts occurred between 1945 and 2003. Some 60% of conflicts started before 1990, and the other 40% started after 1990, thereby making the last decade of the 20th century the decade with the most ethnic conflicts. Since 1955, nearly

50 ethnic groups have been targeted in campaigns of genocides and ethnic cleansings that killed between 13 million and 20 million civilians (Marshall & Gurr, 2005). These civil wars, mass murders, and violent campaigns led to more than 14 million internationally recognized refugees and about 17 million internally displaced people (U.S. Committee for Refugees and Immigrants, 2008). Today, most ethnic conflicts occur in sub-Saharan Africa and Asia.

Many of these conflicts are protracted conflicts, meaning that they have lasted 10 years or more. The Sudanese civil war between the Arab-Muslim north and the Christian-Animist-African south, for example, is the longest and deadliest civil war in the second half of the 20th century. Most ethnic conflicts do not meet the threshold of wars (1,000 or more battle-related deaths in a year). Low-level rebellions, minor armed conflicts (at least 25 battle-related deaths per year), terror campaigns, and large-scale protest movements with occasional violence are more common. Patterns of escalation and de-escalation are typical scenarios. The Sri Lankan civil war, for instance, started in the early 1980s between the Sinhalese government and Tamil rebel groups and ended in 2009 with the defeat of the Tamils. During the conflict, high and low levels of intensity alternated; the conflicting parties negotiated various cease-fires and peace agreements, followed by insurgencies and high levels of violence and death.

The Clash of Civilizations

The idea that cultural differences lead to violent behavior of political actors is also the foundation of Samuel Huntington's famous *Foreign Affairs* article "The Clash of Civilizations?" (1993) and his subsequent book *The Clash of Civilizations and the Remaking of World Order* (1996). From the premise that sameness leads to peaceful relations whereas difference produces disorder and conflict, Huntington argues that cultural and ethnic differences between "civilizations," namely, states or groups of states that distinguish themselves by cultural traits, will lead to conflict. These cultural differences are first and foremost religious in nature, although linguistic and geographic proximity also play a role. He identifies the following as major civilizations:

- Western civilization (western and central Europe, North America, and Australia)
- Latin American civilization (Central and South America)
- Slavic Orthodox civilization (former Soviet Union states [excluding Central Asia], former Yugoslavia [excluding Slovenia and Croatia], and eastern Europe)
- Buddhist civilization (Asian states, including Thailand, Sri Lanka, Laos, Cambodia, Mongolia, and Myanmar/Burma)
- Confucian civilization (China and the Chinese diaspora, North and South Korea, Singapore, Taiwan, and Vietnam)
- Hindu civilization (India and the Indian diaspora, Nepal)

- Japanese civilization
- Islamic civilization (Middle East, North Africa, Central Asia, Indonesia, Malaysia, and Brunei)
- (Sub Saharan) African civilization (southern, central, and eastern Africa)

Some civilizations overlap or are categorized into sub-civilizations (e.g., Western civilization is divided into the European and North American categories; Islamic civilization into Arab, Persian, Turkish, and Indonesian subdivisions). Turkey, Ethiopia, Haiti, and Israel are excluded from this classification of civilizations and designated as so-called *lone countries*.

The idea of classifying the world into civilizations is not entirely new. The British historian Arnold J. Toynbee (1960) concluded, in his book *A Study of History*, that the world consists of 21 civilizations. The German historian and philosopher Oswald Spengler (1918/1991) divides the world into eight cultures in his book *The Decline of the West* and follows a pattern very similar to Huntington's divisions (excluding most of Africa). The term *clash of civilizations* was crafted by British scholar Bernard Lewis (1990) in his article titled "The Roots of Muslim Rage," in which he describes the rivalry between Islam and the Judeo-Christian heritage. Huntington's work endorses Lewis's hypothesis of the clash of civilizations and expands the theory to the global scale.

According to Huntington (1996), future conflicts or "clashes" will happen between these civilizations, either on the local and regional level (what he calls "fault line conflicts") or on the global level between major states of different civilizations ("core state conflicts"). He points out that these conflicts will be mostly between the Western civilization, which currently enjoys hegemonic status, and major challengers, namely the Confucian and Islamic civilizations. East Asia, and above all China, threatens the West mostly because of rapid economic growth, and the rise of fundamentalism in the Islamic world challenges Western values such as liberal democracy and human rights. Huntington sees a potential alignment of these two "challenger civilizations" as both have a history of conflict with the West. In addition, so-called swing civilizations, namely Russia, India, and Japan, who might favor either the West or the challengers, further destabilize the world because their affiliations are unclear, but their power is extensive enough to bring about major changes.

Huntington's (1993, 1996) main prediction is that future conflicts will be fought between Muslims and non-Muslims. Conflicts along boundaries between Muslims and non-Muslims, as in the Philippines, Kashmir, Chechnya, Kosovo, Bosnia, Sudan, Nigeria, and Palestine, are seen as proof that "Islam has bloody borders" (Huntington, 1993, p. 35). Historical clashes of Christians and Muslims dating back to the Middle Ages and the fact that both Islam and Christianity are absolute, universalist religions with a mission to spread their faith are portrayed

as the reasons for current and future problems between the Christian (Western) and Islamic civilizations. The attacks of September 11, 2001, and the subsequent events in Afghanistan and Iraq have been interpreted as proof of Huntington's predictions.

Critics, however, point out that empirical evidence does not support Huntington's thesis. Empirical studies find no increase in the frequency of intercivilizational conflicts and show that state interactions across civilizational divides are not more prone to conflict. Similarly, Huntington's (1996) "kin-country syndrome" (p. 272), namely, the idea that in case of a war, people from the same civilization will support others with the same cultural identity, cannot be empirically established. Scholars have thus disproved major aspects of the theory of the clash of civilizations (Chiozza, 2002; Fox, 2002; Tuscisny, 2004).

Others point to the fact that clear cultural boundaries do not exist in reality. Why separate Japan from China? Why not separate Vietnam from China? Why not distinguish between Catholic and Protestant states in the West? Ideological and philosophical differences, paired with political and economic discrepancies, are the most important factors influencing the likelihood of conflict (Berman, 2003). In turn, ideological and political values such as democratic governance and the rule of law are more easily transmitted than Huntington suggests (Ajami, 1993). Many non-Western states have become democratic over the past decades, and the European Union has expanded beyond western Europe. Cultural perceptions play a secondary role. In addition, many argue that nation-states will remain the major players in international politics. Most conflicts will be fought between states of the same civilization or, more likely, within states (Gray, 1998; Hunter, 1998; Walt, 1997).

Although Huntington's thesis has its merits and has led to considerable scholarly debate, it cannot be empirically proven and has major flaws. Huntington's classification of civilizations is difficult to apply to reality. For example, although all states or groups in the Islamic civilization are mainly Muslim, they express very different worldviews (e.g., Bosnians, Indonesians, and Arabs face completely different economic, social, and political circumstances). In addition, most conflicts today are not conflicts between civilizations but rather conflicts within civilizations and within states. The major causes for conflict today are not cultural differences but economic and political problems, ideological disagreements, and discrimination. It is safe to say that most political scientists today have serious doubts about major elements of Huntington's thesis.

Conclusion

Cultural differences and ethnic conflicts are important issues shaping international politics. Because cultural affiliations and ethnic identity are particularly strong factors

shaping group relations, these conflicts have led to tremendous human suffering and are a significant threat to international security. Instability, refugee flows, spillover effects, and other international consequences guarantee that ethnic conflict remains an issue on the international political agenda. However, it is not the cultural differences per se that lead to conflict but political, ideological, and economic goals of international actors, regardless of whether these actors are states, ethnic groups, or "civilizations." Given the complexity of ethnic and cultural conflicts, there is no "silver bullet solution" to solving related issues.

References and Further Readings

- Ajami, F. (1993). The summoning. *Foreign Affairs*, 72(4), 2-9.
- Anderson, B. (1991). *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.
- Berman, P. (2003). *Terror and liberalism*. New York: W. W. Norton.
- Brass, P. (Ed.). (1985). *Ethnic groups and the state*. London: Croom Helm.
- Brown, M. (2001a). The causes of internal conflict: An overview. In M. Brown (Ed.), *Nationalism and ethnic conflict* (pp. 3-25). Cambridge: MIT Press.
- Brown, M. (2001b). Ethnic and internal conflicts: Causes and implications. In C. A. Crocker, F. O. Hampson, & P. Aall (Eds.), *Turbulent peace: The challenges of managing international conflict* (pp. 209-226). Washington, DC: U.S. Institute of Peace Press.
- Brubaker, R. (1995, Spring). National minorities, nationalizing states, and external national homelands in the new Europe. *Daedalus*, 124, 107-132.
- Chiozza, G. (2002). Is there a clash of civilizations? Evidence from patterns of international involvement, 1946-97. *Journal of Peace Research*, 39(6), 711-734.
- Chipman, J. (1993). Managing the politics of parochialism. In M. Brown (Ed.), *Ethnic conflict and international security* (pp. 237-264). Princeton, NJ: Princeton University Press.
- Dominguez, V. R. (1989). *People as subject, people as object: Selfhood and peoplehood in contemporary Israel*. Madison: University of Wisconsin Press.
- Esman, M. J. (1994). *Ethnic politics*. Ithaca, NY: Cornell University Press.
- Fearon, J., & Laitin, D. D. (2003, February). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1), 75-90.
- Fox, J. (2002). Ethnic minorities and the clash of civilizations: A quantitative analysis of Huntington's thesis. *British Journal of Political Science*, 32, 415-434.
- Geertz, C. (1973). The integrative revolution: Primordial sentiments and civil politics in new states. In C. Geertz (Ed.), *The interpretation of cultures* (pp. 255-311). New York: Basic Books.
- Glazer, N., & Moynihan, D. P. (1975). *Ethnicity: Theory and experience*. Cambridge, MA: Harvard University Press.
- Gray, J. (1998). Global utopias and clashing civilizations: Misunderstanding the prosperity. *International Affairs*, 74, 149-164.

- Gurr, T. R. (1993). *Minorities at risk: A global view of ethno-political conflicts*. Washington, DC: U.S. Institute of Peace Press.
- Gurr, T. R., & Harff, B. (2003). *Ethnic conflict and world politics* (2nd ed.). Boulder, CO: Westview.
- Harbom, L., & Wallensteen, P. (2007). Armed conflict, 1989–2006. *Journal of Peace Research*, 44(5), 623–643.
- Horowitz, D. L. (1985). *Ethnic groups in conflict*. Berkeley: University of California Press.
- Horowitz, D. L. (2001). *The deadly ethnic riot*. Berkeley: University of California Press.
- Hunter, S. (1998). *The future of Islam and the West: Clash of civilizations or peaceful coexistence?* Westport, CT: Praeger.
- Huntington, S. P. (1993, Summer). The clash of civilizations? *Foreign Affairs*, 72, 22–49.
- Huntington, S. P. (1996). *The clash of civilizations and the remaking of world order*. New York: Simon & Schuster.
- Isaacs, H. (1975). *Idols of the tribe: Group identity and political change*. New York: Harper & Row.
- Kuperman, A. J. (2004, Spring). Humanitarian hazard: Revisiting doctrines of intervention. *Harvard International Review*, 26(1), 64–68.
- Laitin, D. (1986). *Hegemony and culture* (2nd ed.). Chicago: University of Chicago Press.
- Lake, D. A., & Rothchild, D. (Eds.). (1998). *The international spread of ethnic conflict: Fear, diffusion, and escalation*. Princeton, NJ: Princeton University Press.
- Lewis, B. (1990, September). The roots of Muslim rage. *The Atlantic Monthly*. Retrieved May 20, 2009, from http://www.theatlantic.com/doc/199009/muslim_rage
- Marshall, M. G., & Gurr, T. R. (2005). *Peace and conflict*. Retrieved May 30, 2010, from http://www.cidcm.umd.edu/publications/papers/peace_and_conflict_2005.pdf
- Noel, D. L. (1968). A theory of the origin of ethnic stratification. *Social Problems*, 16(2), 157–172.
- Posen, B. A. (1993). The security dilemma and ethnic conflict. In M. Brown (Ed.), *Ethnic conflict and international security* (pp. 103–124). Princeton, NJ: Princeton University Press.
- Ross, M. H. (2001). Psychocultural interpretations and dramas: Identity dynamics in ethnic conflict. *Political Psychology*, 22, 157–178.
- Smith, A. D. (1986). *The ethnic origins of nations*. Oxford, UK: Blackwell.
- Smith, A. D. (1991). *National identity*. London: Penguin.
- Spengler, O. (1991). *The decline of the West* (A. Helps & H. Werner, Eds.; C. F. Atkinson, Trans.). New York: Oxford University Press. (Original work published 1918)
- Toynbee, A. J. (1960). *A study of history*. London: Oxford University Press.
- Tusicsisny, A. (2004). Civilizational conflicts: More frequent, longer, and bloodier? *Journal of Peace Research*, 41(4), 485–498.
- U.S. Central Intelligence Agency. (n.d.). *World factbook: Albania*. Retrieved April 29, 2010, from https://www.cia.gov/library/publications/the_world_factbook/geos/al.html
- U.S. Committee for Refugees and Immigrants. (2008). *World Refugee Survey, 2008*. Retrieved April 29, 2010, from <http://www.refugees.org/article.aspx?id=2114&subm=179&area>About%20Refugees>
- Van Evera, S. (1994). Hypotheses on nationalism and war. *International Security*, 18(4), 5–39.
- Volkan, V. (1997). *Bloodlines: From ethnic pride to ethnic terrorism*. New York: Farrar, Straus & Giroux.
- Walker, C. (1994). *Ethnonationalism: The quest for understanding*. Princeton, NJ: Princeton University Press.
- Walt, S. N. (1997). Building up new bogeymen. *Foreign Policy*, 106, 177–189.
- Wolff, S. (2006). *Ethnic conflict: A global perspective*. Oxford, UK: Oxford University Press.

18

COMPARATIVE POLITICAL PARTIES

Systems and Organizations

ALEXANDRA COLE

California State University, Northridge

Political parties and party systems are of interest to the scholar of comparative politics because they are constantly in flux. A common understanding of the political party, according to Leon D. Epstein (1967), is of a group that “seeks to elect governmental officeholders under a given label” (p. 9). *Party systems* are described by the number of parties within a given country during a given time, along with their “internal structures, their ideologies, their respective sizes, alliances, and types of opposition” (Duverger, 1972, p. 18). Party systems can have as few as one major political party, or may have many political parties. Elections are the venue in which competition for government office takes place. Elections bring changes in the policies advocated by parties, the seats held by political parties, and of course the composition of government. It is important to note that political parties do not make changes in a vacuum; change often comes in anticipation of, or in reaction to, changes that other political parties in the system make. This makes the party system a system of interaction between political parties (Sartori, 1976).

This chapter will examine political party systems in a comparative context. We will begin with a broad discussion of political parties. The difference in number, type, and ideology of political parties across different party systems has much to do with the political development

of a polity. Though there may be similarities in the ideologies of political parties in different systems, the parties may behave differently because of the dynamics within their own systems. Political parties that would never work together in one system because of ideological differences may be coalition partners in another system. This may have to do with the electoral rules of a system and the prospects for formation of a coalition government, or it may have to do with attitudes toward the political system in general. We will see that ideological considerations often have less to do in explaining the behavior of a political party compared with the potential for policy outputs. We end the chapter with a discussion of whether political parties are in decline, and the potential effects of such a decline.

The Formation of Political Parties

We understand *political parties* as organizations that regularly compete for public office in that they put forth candidates for election (Sartori, 1976). The formation of political parties is generally associated with the extension of suffrage and the development of representative government (Duverger, 1972). Joseph LaPalombara and

Myron Weiner (1966) suggest that political parties are endemic to “modern and modernizing political systems”: A political party will emerge once a “political system reaches a certain degree of complexity, or whenever the notion of political power comes to include the idea that the mass public must participate” (p. 3). These definitions seem to place political parties as 20th-century phenomena, although Seymour Martin Lipset and Stein Rokkan (1967) theorized that many of the political parties that existed at the end of the 20th century were based in part on earlier political conflicts dating back to at least the 19th century if not earlier. Specifically, these conflicts ranged from national revolutions to the political aftermath of the Industrial Revolution. From these events arose cleavages, or divisions, within societies that gave rise to political groupings. Lipset and Rokkan suggest that there are hierarchies in cleavages within systems and over time, which helps to explain the differences in political groupings across countries. It is important to note that the cleavages themselves may also change, or even lose relevance.

The most common cleavages can be classified as territorial and functional. *Territorial* cleavages arise when there is conflict between the central nation building culture and that of the periphery. Peripheral cultures are those differing in ethnicity, language, or religion from the center of the nation (Lipset & Rokkan, 1967), which often places the peripheral culture in the position of subject culture. *Functional* cleavages can be interest specific or ideological. Interest-specific cleavages are those that reflect conflict over resources (Lipset & Rokkan, 1967), whereas ideological cleavages often reflect differing worldviews. For example, in national revolutions the conflict between the nation-building center and periphery represents a territorial cleavage, whereas the often accompanying conflict between church and state results in a functional cleavage of religious versus secular worldviews. Industrial revolutions, which can pit the interests of industrialists against those of landed interests, can be territorial, whereas the conflict between owners and workers is functional in nature. Together, these constitute the four critical lines of cleavage that Lipset and Rokkan (1967) suggest explain the variance in many modern competitive party systems.

The transformation of a conflict into a cleavage and then a political grouping, or party, takes place only after particular thresholds are crossed in the development of a nation-state. These thresholds are (a) *legitimation*: Is there recognition of the right of protest? (b) *incorporation*: Are supporters of a movement given political citizenship rights? (c) *representation*: Can the new movement exist on its own, or must it join with older movements? and (d) *majority power*: Are there checks and balances against numerical majority rule? (Lipset & Rokkan, 1967). The first two thresholds specifically influence the development of a political party, whereas the latter two are related to the

growth and development of the party system. As the first two thresholds occur roughly at the same time as the extension of mass suffrage, the contours of the party system are set relatively early in the life of the nation-state; thus the observation from Lipset and Rokkan (1967) that the “crucial differences among the party systems emerged in the early phases of competitive politics before the final phase of mass mobilization” (p. 114).

Lipset and Rokkan suggest that much of the development of party systems in Western Europe followed this model. Although political parties may differ, overall one may speak of *party families*, in which ideological tendencies are reflected by parties across different states. Beyme (1985) suggests that parties based on ideological principles have had more success in establishing themselves in western Europe than have parties based on specific conflicts. The earliest modern political parties, *liberal parties*, established themselves as supporters of representative democracy and constitutionalism. This involved the recognition of individual rights and the preservation of individual property. Within liberalism was a secondary, more radical branch that supported direct rule by the people, meaning the extension of suffrage to the masses. Liberal parties first emerged in England in the 1700s and in France after 1830. In many cases, they were the first parties to form in representative democracies. Over time, they have become smaller parties in many political systems. Though smaller, liberal parties such as the Free Democratic Party of Germany often govern as part of a ruling coalition.

Conservative parties developed alongside liberal parties, often as a response to liberal parties. The conservative ideal is the preservation of historical continuity, with a belief in the divine, valuation of traditional forms of life, and recognition of private property and freedom (Beyme, 1985). Conservative parties serve as a buffer to liberal parties because conservative parties oppose rapid change, which is presumed to threaten the social order. Conservative parties tend to be in more secularized political systems. England's Conservative Party is one of the most well-known conservative parties.

Both liberal and conservative parties established themselves while representative governments formed, primarily reacting to the conflicts within a particular political system. The constituencies for liberal and conservative parties were necessarily small; not all the inhabitants of a democratizing polity had the franchise, and liberal and conservative parties tended to represent the upper classes. In contrast, *socialist* parties were the first parties to form outside representative bodies of government, often at the same time in which the franchise was extended to all. According to Beyme (1985), socialists were aiming for revolutionary reversal rather than maintenance of the “principles of 1789”: liberty, equality, and fraternity. Socialist parties tended to mobilize the newly enfranchised working classes. Mostly worker based and highly unionized, socialist parties sought better conditions

for workers and demanded more state intervention in the economy. The constituency for socialist parties has largely stayed the same, as can be seen in cases such as the French Socialist Party. After World War II, socialist parties largely abandoned calls for full state intervention in the economy, instead focusing on implementation of stronger control mechanisms in the economy.

Communist parties are often linked to socialist parties in terms of their advocacy for workers' rights and state-controlled economies. However, their historical difference with socialists dates to World War I, when some believed socialist parties were not strong enough in their opposition to the war (Beyme, 1985). The antiwar attitude unified many communist parties across European nations and led to the formation of a communist movement. This communist movement was much more ideological than other parties were, resulting in an outlook that was more international than national. Within older democratic systems, such as in France and Italy, communist parties have retained some influence, although their significance has waned in many other countries. In post-Communist states, the personnel of former Communist parties has remained in politics, although the parties themselves have undergone some changes, especially name changes. The former East German Communist Party became the Party of Democratic Socialism after unification and more recently transformed itself into the Left Party.

Christian democratic parties also originated before World War I. As some nations democratized, in some of the more religious nations, the established church found itself at odds with the secularizing tendencies of liberal reformers. Christian political parties were "generally formed as a defense counter reaction to liberal or secular legislations by which ardent believers felt threatened" (Beyme, 1985, p. 81). Only after the excesses of the national socialist era did established churches begin to realize the importance of democratic forms of government (Beyme, 1985). Accordingly, Christian democratic parties became popular in the post-World War era because their ideological orientation tended toward the center, especially economically, with a religious focus on moral issues.

Smaller parties such as *agrarian* or *regional* or *ethnic* parties are not as widespread as the previous party families are, owing to the particular historical circumstances of their formation. According to Beyme (1985), "Agrarian parties only emerged in countries where the towns were still relatively small during the period of the extension of the franchise and the rural population was strong enough to stand up to the major landowners" (pp. 112–113). The most successful agrarian parties have been in some of the Scandinavian countries. Agrarian parties tend to be to the right on political and social issues and were at their peak during the interwar period. Most have been subsumed by other parties.

Regional or ethnic parties also reflect a carryover from the nation-building process, and representation issues to

this day still divide the regional or ethnic minorities from the larger society. Most advocate for self-determination or autonomy in their own affairs. The most famous of these are the Basque Nationalist Party, in Spain, and the Scottish Nationalist Party, in Scotland. These parties advocate for the independence of the Basque and Scottish peoples from Spain and the United Kingdom, respectively.

The second half of the 20th century witnessed the emergence of political parties not traceable to the cleavages of national or industrial revolutions. The cleavages that would be politically relevant would be based less on redistributive issues and more on what Ronald Inglehart (1977) refers to as postmaterialist, or quality-of-life, issues. The emergence of *ecological* parties such as green parties marked the transformation of this cleavage into a political grouping. As opposed to other political groupings, green parties seemed to lack a particular social base of support and represented issues such as the environment, nuclear power, human rights, and democratic representation. These issues could not easily be placed within a traditional left–right understanding of politics and soon came to be known as issues of the *New Left*. Not surprisingly, the social and political developments that led to a New Left cleavage would also lead to a *New Right*, in which law and order, patriotism, and personal morality issues were similarly difficult to place within a traditional left–right understanding of political parties. Parties of the New Right are also referred to as right wing extremist parties or parties of the far right, but they share an emphasis on the above issues. The most famous of these has been the Front National of France.

Party Organizations

Political parties tend to differ in their organization on the basis of three factors: competition, institutionalization, and resource factors (Ware, 1996). *Competition* refers to ideological differences, as well as the way in which the party was formed. *Institutionalization* refers to the power relations both between parties and within parties. *Resources* refers to how the party perpetuates itself. Each of these factors is affected by the specific time in which a party emerges. Different organizational structures are thus related to specific social and political developments in the modern democratic state. Organizational differences become apparent when one observes what Richard Katz and Peter Mair (1993) term the different *faces* or responsibilities of political parties. The three faces are the party on the ground, the party in central office, and the party in public office. The *party on the ground* refers to the political party as represented by the electorate, or the voters a party can reasonably rely on to vote for it. The *party in central office* refers to the membership aspect of a political party, or those who actively participate within the party with respect to policy formation, recruitment of members, and campaign planning. The *party in public*

office refers to the members of a political party who serve as elected representatives.

The earliest political parties, known as *cadre*, or *elite*, parties, predated mass suffrage. Elite parties were small parties that largely reflected the interests of the elite classes. Because suffrage was limited, political representation of those who could vote was also limited—essentially narrow constituency groups, often locally based. The party in the electorate was indistinguishable from the party in public office because those elected to office came from local constituencies and directly represented voters. Organizationally, elite parties were not complex. A clear correspondence between voters and representatives existed, and local interests were well represented by the local representatives elected to national legislative bodies. Katz and Mair (2002) summarize the elite party as follows:

a small party on the ground in each constituency able to provide its own resources, close and locally based ties between the individual members of the party in public office and the individual parties on the ground, weak or entirely absent party in central office. (p. 116)

The extension of mass suffrage, well under way by the middle of the 19th century, not only coincided with emerging political ideologies representative of the interests of the working classes, but also led to the formation of *mass* parties. As Peter Mair (1990) writes, “The extension of [suffrage] incorporated the mass of the citizenry into the political system; mass parties mobilize and integrate these citizens and inculcate a set of enduring political identities” (p. 4). Mass parties are parties with a focus on national issues and thus on winning national representation. Unlike elite parties, mass parties depend on the maintenance of high levels of party membership. This is partly because of the need to attract votes and gain political representation but also because of the need to attract resources, most obviously financial ones. Elite parties, because of their small size, can rely on the support of wealthy backers, whereas mass parties need to appeal to large numbers of newly enfranchised working-class voters because mass parties need the dues of their members to remain financially solvent. As a consequence, the organization of mass parties is much more complex than that of elite parties. Katz and Mair (2002) describe the case of the mass party as follows:

The party in central office provides support for the expansion of the party on the ground and central coordination for its activities, while the party on the ground provides the resources that are necessary for the existence and success of the party in central office. (p. 117)

The necessity of coordinating activities between the party on the ground and the party in central office led to an increase in the importance of professional staff members. Adding to the necessity of strong organization is the importance of coordination with the mass party’s elected

representatives, or the party in public office. Not only are elected officials answerable to the constituencies that vote them into office, but they also must answer to the party in central office, responsible for the electoral activities of the party. With mass parties, the importance of party organization is evident in the need to coordinate between the three faces of the party, on the ground, in public office, and in central office.

Understandably, elite parties, which represented a narrow constituency, would lose relevance in democratic societies when the franchise was extended. According to Otto Kirchheimer (1966), mass parties themselves would also begin to fade after World War II as the societal cleavages that Lipset and Rokkan (1967) described lost some of their relevance. The political and economic development of the modern state made distinctions based solely on class or denomination less divisive. Along with social and political changes within the electorate, political parties themselves changed, now with an increase in emphasis on winning elections and gaining seats in national legislatures. To win elections and gain seats, political parties needed to broaden their appeal past the narrow clienteles of the elite parties, or even the specific class-based focus of the mass parties. Some political parties developed a *catchall* approach, in which the aim was to catch all categories of voters, not just traditional constituencies based on societal cleavages. Ideological considerations were less important than not alienating a particular constituency group. Catchall parties would appeal to the median voter in society as opposed to a specific section of the electorate.

Organizationally, the catchall party differed from the mass party model on many fronts. New forms of media, such as television, made the activism associated with mass parties less of a crucial element for catchall parties. Rather than extend the effort involved in appealing to voters through personal contact and activist organization, political parties found that they could appeal to more voters through the media. It is important to note that the reach of the media also meant that catchall parties would emphasize the recruitment of party leaders who could appeal to the widest swath of voters. For the first time, the party in public office and the party in central office would be the most important faces of the party, as opposed to the party in the electorate and party in public office faces of elite parties and the party in the electorate and party in central office faces of mass parties. Catchall parties were formed specifically to win elections, and the way to do so was under the direction of a central office charged with the responsibility of running election campaigns and choosing the best representatives, from the point of view of the party, to stand for them.

Changes in political party organization echo changes in society and in politics. Modern political parties have placed more importance on winning elections, even though the importance of party membership has decreased. Catchall parties found they could win elections by appealing to the widest possible bloc of voters. At the same time,

voters have become less likely to identify with a specific political party and more likely to shift their allegiances from election to election. With shifting voters comprising an ever larger proportion of the electorate, political parties are less likely to rely on voters for resources. Katz and Mair (1995) suggest a new type of party has emerged that has adapted to these realities: the *cartel* party. Cartel parties are characterized as comprising professional politicians whose main source of support is actually the state and public sources of financing for political parties. Although political parties may compete against each other for votes and seats, all implicitly understand that their survival depends on maintenance of office rather than ideological battles. Katz and Mair write that “as politicians pursue long term careers, they come to regard their political opponents as fellow professionals” (p. 23). The party in public office is the most important facet because elected officials both attract votes and make sure that sources of public funding remain in place for themselves.

Party Systems

Both the competition and the prospects for cooperation between political parties in an electoral system constitute a *party system*. Party systems may differ on the basis of the types of parties within a system (both ideological and organizational) and the number of parties within a system. Which parties are included as part of a party system is decided on the basis of what Giovanni Sartori (1976) terms *coalition potential* and *blackmail potential*. *Coalition potential* refers to whether a party can be considered an acceptable coalition partner in order to control government. This definition does not imply that a party has to be in government to be considered a party of the party system, but rather that it has the potential to be part of a governing coalition. *Blackmail potential* refers to whether a party can affect the tactics of party competition of the parties that have coalition potential. This definition does not imply that a party must be part of a coalition, or have any chance of being part of a coalition, but that it can influence the political parties that do have coalition potential.

Early observers of political party systems such as Maurice Duverger (1951/1954) held that the number of parties within a system should be the main criterion for defining a party system. The number of political parties within a party system is largely dependent on the specific election rules of that political system. By *election rules*, Duverger meant the barriers to representation, or what percentage of the vote a party must secure in order to be represented in the national legislative body. The main difference is between systems following majoritarian representation rules and those following a proportional representation rule. A *majoritarian* system is one in which a party (or its candidate) must secure more than 50% of the popular vote. This type of system is also commonly referred to as a *first past the post* system, with *the post* referring to 50% of the vote: The party

that first gets 50% of the vote gets representation. Majoritarian systems tend to limit the number of parties that compete in these systems because the parties must necessarily appeal to the widest range of voters. Were a party to appeal only to one or the other side of the political spectrum, it would only have the votes of a minority of voters. Thus, Duverger held that majoritarian systems tended to be two-party systems, with the parties themselves more moderate in their political ideologies because of the necessity of having to appeal to a wider group of voters. Two-party systems tend not to have cooperation between the major parties, given that one party necessarily has a majority of the seats in the legislative branch and thus does not need the opposing party to form policies. The United States is one of the most notable two-party systems.

Political systems that follow *proportional* representation tend to have a greater number of political parties because parties will win seats in the national legislature based on their percentage of the popular vote. Some systems, such as Germany, have instituted minimum-vote percentages, which lower the probability of extremist parties' gaining representation because parties must win at least a specific percentage of the vote to gain seats. In political systems in which there is a minimum-vote threshold, the mean number of political parties tends to be lower than in systems in which there is no minimum-vote threshold. Even so, there is no standard number of parties within a multiparty system. *Two and a half party systems* are systems that have three parties, with the third party much smaller than the other two. Australia and Canada are notable two-and-a-half-party systems. The third party tends to alternate as a coalition partner between the larger parties, although a *grand coalition* between the larger parties is not unheard of. In a *system with one large party and several smaller parties*, the larger party tends to be in power for long periods, with a coalition of the remaining parties necessary to unseat the larger party. The party systems of Norway and Sweden exemplify this type of system. *Systems with two larger parties and several smaller parties* necessitate the formation of coalitions between the larger parties and some, or several, of the smaller parties. In this case, a grand coalition between the larger parties is unlikely because of ideological distance. Israel can be said to illustrate this type of system. Finally, *even multiparty systems* consist of a broad category of multiparty systems that can range from systems in which there is complete cooperation between the parties to polarized and volatile party systems (Ware, 1996). Italy in the postwar era was long the main example of this type of system. The number of parties in a system is not sufficient in itself to describe the nature of the party system; the nature of party competition is an important component in the classification of party systems.

Following Duverger, Sartori (1976) suggested that ideological distance, as well as party fragmentation, determine the nature of party competition within party systems. *Ideological distance* is defined as the “overall spread of the ideological spectrum of any given polity” (p. 126). At the

time of Sartori's work, this primarily referred to parties aligned along a traditional left–right continuum. Ideological distance also refers to the attitudes of political parties toward the state, as well as toward other parties within the system. Thus, political parties may be close ideologically but differ in how they perceive themselves in relation to the state and to each other. Extremist parties, although having similar ideological tendencies as parties of the left or the right, may nevertheless be considered ideologically distant if their ideology incorporates antisystem tendencies.

Party fragmentation incorporates the number of parties within the system and whether any of the parties “approaches the absolute majority point” (Sartori, 1976, p. 124). The more parties within a party system, the more likely it is that a party system will be fragmented, especially if there is a large ideological distance between the parties. A large number of parties plus a large ideological distance between the parties can result in what Sartori refers to as a *centrifugal system*. A centrifugal system of party competition is one in which most parties exist at the extremes of the system, with a vacuum in the political center of the system. Conversely, a *centripetal system* of party competition displays a pull to the center for the political parties and is much more likely in systems with smaller numbers of political parties.

When party fragmentation is taken into account, the classification of party systems becomes more complex. Two-party systems tend to have lower levels of party fragmentation in general, although Sartori (1976) warns not to assume that they “always work.” Rather, Sartori suggests that the “centripetal mechanics of twopartism creates consensus” (p. 191). In a *predominant party system*, in which one party has the majority of votes although other parties are represented in the system, as long as the predominant party retains the majority of seats, party fragmentation is also low. Japan in the postwar era is an example of this. However, should the predominant party lose an outright majority, the system may change into one in which party fragmentation is higher.

Moderate multipartism describes party systems of about three to five parties, with moderate levels of party fragmentation and centripetal tendencies within the system. That is to say, the parties are pulled toward the center. In contrast, *polarized multipartism* describes party systems with about three to five parties and centrifugal tendencies. The center is weak in such systems, which are more likely to suffer instability. Societies that have deep cleavages within them, and many parties to represent these cleavages, tend to display *segmented multipartism*. Although the political tendencies are centrifugal, or a lack of center exists, such systems can endure through institutional design. The best examples of segmented multipartism through institutional design are the consociational democracies such as the Netherlands, described by Arend Lijphart (1969).

Party systems are defined through the prospects for party competition and party cooperation in a political system.

Party competition is a straightforward concept; parties compete with each other for votes and seats in a legislature. Party cooperation, however, has a slightly different focus in that it describes how parties interact with each other after receiving votes and seats in a national legislature. Levels of party cooperation are determined by the willingness with which parties will go into coalitions in the formation of a government. In a predominant party system, or a two-party system, party cooperation is not a necessity as one party has a majority of the votes and seats. However, in party systems with more than two parties, coalitions are a necessity in order to form a government because there is no clear majority party. The extent to which political parties can cooperate in the formation of a ruling coalition is dependent on various factors and is the subject of the following section.

Formation of Government

One of the functions separating political parties from other interest groups in society is the translation of issue preferences into policy. Ian Budge and Hans Keman (1990) suggest that contrary to some conceptions of political parties, winning elections is not the most important goal for political parties; formulating policy is. Per Budge and Keman, “explaining the behavior of parties in government is a natural corollary to explaining how they gain the popular support necessary to sustain a governmental role” (p. 2). Without the ability to translate preferences into policy, political parties would not have support from the electorate. Thus the extent to which parties can make or influence policy is a key determinant of their longevity within a political system.

Within majoritarian political systems, the party that receives the most votes forms the government because it controls the most seats in the legislative branch. In multi-party systems, the formation of a government is much more complicated. Budge and Keman (1990) offer a general theory of party government to explain the factors influencing the party coalitions that may form when no one party controls a majority of votes. According to Budge and Keman, this general theory has four assumptions:

1. The party or combination of parties that can win a legislative vote of confidence forms the government.
2. Parties seek to form a government that can survive legislative votes of confidence and most effectively carry through policy.
3. The chief preferences of all democratic parties is to counter threats to the democratic system; where no such threats exist, the chief preference is to carry through differences related to issues along the socialist bourgeois¹ dimension of issue competition; where these two threats do not hold, the preference is to pursue group related preferences.
4. Within parties, factions seek to transform their issue preferences into policies.

The implications of this general theory of party government affect explanations of party behavior, how governing coalitions form and how they change, what governments do in terms of their policy outputs, and how governments come to an end. The formation of governments refers not only to which parties are part of the ruling coalition but also to how government ministries are distributed among the coalition partners. In fact, Budge and Keman (1990) suggest that parties may influence government more through their tenure of specific ministries than through the negotiations that lead political parties into coalition.

The coalition process begins when it becomes apparent that no one party has enough legislative seats to control the government. William H. Riker (1962) suggested that the most obvious coalition to form would be a *minimal winning coalition*, in which there are enough members within the coalition to assure control of government, but no surplus members. This approach explains the behavior of political parties if their main goal is the maintenance of office but does little, according to Budge and Keman (1990), to explain the formation of policy. A policy-based approach to explaining coalition formation may be better at “explaining why governments adopt the kind of policy they do” (p. 19) and ultimately how responsive elected governments are to voters.

In forming coalitions, then, parties take into account the seats held by other parties within the system, as well as the policy positions of all parties within the system. Coalitions tend to form on the initiative of the largest parties within the system (Budge & Keman, 1990). Parties are more likely to enter into coalition with other parties that share the same policy preferences as they do and that have enough legislative seats to form a coalition of at least 50% plus one of all legislative seats. Policy positions are not static; they can change, given historical situations, which implies that particular coalitions are not necessarily a given. Nevertheless, parties will enter into the coalition that they figure will provide the best possibility of implementing their policy preferences, and they do so based on a calculation of policy preference overlap between the parties within a system. The more overlap in terms of policy preference, the more stable over time the coalition will be. Coalitions can form in the absence of policy overlap; in more fragmented systems, prosystem attitudes may be enough to enter into coalition, although these coalitions tend to be the least stable over time. Generally speaking, the smaller the coalition, the more stable the coalition tends to be.

Coalition agreements specify not only which parties will control the government but also which ministries are held by the specific coalition members. Generally, the largest coalition member holds the prime ministry, with other ministries allocated on the basis of the policy interests of the specific coalition members. The most likely scenario is one in which the number of ministries held by a coalition partner reflects the proportion of seats it holds within the coalition (Budge & Keman, 1990). In a broad sense, the policy priorities of political parties will differ by

political families, so the distribution of particular ministries is somewhat predictable from the party family of a coalition partner. For example, an agricultural party would reasonably be expected to retain the ministry of agriculture within a coalition. If there are potential conflicts between coalition members over ministries, the parties will bargain over ministries until the ministries have been allocated to reflect the proportional distribution of seats held by the coalition members.

In their examination of coalition formation and government functioning in 20 states over time, Budge and Keman (1990) noted a correspondence between parties and the policies that governments made, with parties clearly moving “policies in the direction of their own preferences and values” (p. 158). The importance of policy also plays a role in the termination of governing coalitions; when there are policy differences between coalition partners, the termination of a governing coalition is more likely to take place. A single-party government tends to last longer than governing coalitions because of the absence of policy difference. However, the single most important cause of the termination of government is in fact an election. Voters ultimately decide on the longevity of a government. If they do not like the policies of a government, they are more likely than ever before to vote against the parties of a governing coalition. An overall trend of less stable voting patterns among voters is a major factor in this development.

Decline of Parties

Party identification is defined as a long-term psychological identification with a particular political party (Campbell, Converse, Miller, & Stokes, 1960) and has long been one of the most reliable indicators of the individual vote. If a voter identifies strongly with a political party, the voter is likely to vote for that party in an election. Political partisanship is primarily transmitted during childhood; children will imitate their parents in terms of the political parties they identify with, and this identification lasts well into adulthood. Family is not the only agent of socialization; education, occupation, and social networks serve as alternative venues in the transmission of political partisanship. Although party identification can be influenced by social demographic factors, ideological and issue orientations play a role in the identification of a voter with a specific political party. Party identification has a central role in the study of democracies because political parties provide a linkage for the voters with their government. Political parties can serve as an information cue for voters in elections by educating voters on political issues and candidates, and political parties can mobilize voters to vote in elections.

The discovery of a decline in the percentage of citizens identifying with a specific political party at the end of the 20th century has led to some concern. Initially, partisan *realignment* was thought to be taking place because of

fluctuations in elections in the 1970s and 1980s (Dalton, 2008). Voters were not voting consistently for the same parties over election cycles, as they once had. *Partisan realignment* is the conversion, or realignment, of large numbers of voters from one political party to another political party. Public opinion research supported a different argument: Voters were not realigning themselves, but instead were displaying *dealignment*, or an erosion in party loyalties. The evidence for dealignment included a decrease in party loyalty, lower levels of confidence in parties as political institutions, and an increase in the percentage of voters who not only shifted their votes from election to election but also waited longer to make their choices (Dalton, 2008).

The sources of voter dealignment are said to range from a decrease in the importance of sociopolitical cleavages to changes in the mass media and to changes in political parties themselves (Dalton & Wattenberg, 2000b). The *modernization hypothesis* put forth by Ronald Inglehart (1997) suggests that socioeconomic changes after World War II have led to higher levels of education and standards of living, which have led to an erosion of group-based politics based on class. If the cleavages in society that led to the formation of political parties no longer apply, then the relevance of these parties would also seem less applicable. Higher levels of education, coupled with changes in the mass media, also play a role in partisan dealignment. If the mass media have assumed many of the information functions that political parties once performed (Dalton & Wattenberg, 2000b), then it stands to reason that parties would lose some of their relevance. Finally, changes in political parties themselves, such as an increased emphasis on candidates over party ideology, have led more people to vote on the basis of specific issues and candidates, which further decreases the relevance of political parties.

Kay Lawson and Peter Merkl (1988) suggest that major “parties fail when they do not perform the functions they are expected to perform in their own society” (p. 5). The emergence of interest groups, single-issue movements, and different forms of political organization as motors of interest aggregation—one of the primary functions of political parties—serves as further evidence of party decline. In some systems, parties may fade away, while in other systems, new parties based on political movements may emerge. Party decline is not a uniform phenomenon but is influenced by the type of political system in which the party is located. Even so, Lawson and Merkl note that although there is evidence of party decline, the persistence of political parties in general suggests the continued relevance of political parties, although in different ways.

Implications for the Future of Parties and Party Systems

Dalton and Wattenberg (2000a) indicate that political parties have made adaptations in the face of evidence of their

decline. The emergence of cartel parties is one example of party adaptation. Dalton and Wattenberg further suggest that “parties are benefiting themselves (financially and electorally) at the expense of some of the functions that have made them so essential to the democratic process, such as socialization, mobilization and representation” (p. 269). Given that voters have a declining propensity to identify with the same party over time, and an increased propensity to change their identification from election to election, this loosening of the linkage between parties and voters leads to higher volatility within the electorate. More distressing to Dalton and Wattenberg is the possibility that parties may become less responsive to voters because of the decrease in ties to voters: “If organizational maintenance becomes a party’s primary goal, democracy will inevitably suffer” (p. 270).

Even so, the emergence of cartel parties, which may block the emergence of new parties as challengers within the political system, may not stem all democratic opposition or democratic representation. That is, cartel parties may limit competition among themselves but are unable to limit political opposition and challenges from outside the cartel (Katz & Mair, 1995). Referring again to Lawson and Merkl (1988), the emergence of social movements and single-issue groups suggests that interest articulation is alive and well in democratic societies. Although it is the case that there are lower levels of partisan identification in the electorate and lower levels of trust in political parties, some suggest that these developments reflect another stage in democracy itself (Dalton, 2008). Higher levels of education and cognitive mobilization have led to more politically sophisticated citizens who are even more likely to participate politically, although not through political party mechanisms.

Throughout much of the literature on political parties and political party systems, a common theme has been that of change. Over the stages of democratic development, we see that there have not been single types of parties in specific periods but rather many different types of political parties over extended periods. The organizational forms of parties tend to reflect changes within the broader system, be they ideological shifts, organizational shifts, or even shifts in the competitive framework. An example of this was seen in the transition from mass parties to catchall parties based on technological and social changes. Ideological shifts also take place over time; a party such as the Austrian Freedom Party was considered an example of a liberal party in the 1960s and 1970s but by the end of the 1980s was considered to be more an example of the New Right (Cole, 2005). As political parties remain part of the democratic framework, they must necessarily adapt to account for a greater proportion of the electorate with lower levels of partisanship but higher levels of political sophistication. This need may imply further changes in party organizations, or it may reflect different ideological orientations. It may also mean the fading of

parties from party systems or the inclusion of new parties in party systems. Much as the formation of political parties and party systems reflected political conditions at their founding, transformations of parties and party systems reflect political conditions. This is perhaps the only constant in an area of study that is based on constant change.

Note

1. "Socialist bourgeois dimension" refers to politics along a left right continuum, with the continuum reflecting socioeconomic policy preferences. Socialist policies exist on the left side of the continuum whereas bourgeois, or capitalist, policies exist on the right side of the continuum. Budge and Keman (1990) suggest that the "clash between traditional and innovative societal values" can also be placed on this continuum (p. 20).

References and Further Readings

- Beyma, K. von. (1985). *Political parties in Western democracies* (E. Martin, Trans.). Aldershot, UK: Gower.
- Budge, I., & Keman, H. (1990). *Parties and democracy: Coalition formation and government functioning in twenty states*. Oxford, UK: Oxford University Press.
- Campbell, A., Converse, P., Miller, W., & Stokes, D. (1960). *The American voter*. New York: Wiley.
- Cole, A. (2005). Old right or new right? The ideological positioning of parties of the far right. *European Journal of Political Research*, 44(2), 203–230.
- Dalton, R. J. (2008). *Citizen politics* (5th ed.). Washington, DC: CQ Press.
- Dalton, R. J., & Wattenberg, M. P. (2000a). Partisan change and the democratic process. In R. J. Dalton & M. P. Wattenberg (Eds.), *Parties without partisans: Political change in advanced industrial democracies* (pp. 261–284). Oxford, UK: Oxford University Press.
- Dalton, R. J., & Wattenberg, M. P. (2000b). Unthinkable democracy: Political change in advanced industrial democracies. In R. J. Dalton & M. P. Wattenberg (Eds.), *Parties without partisans: Political change in advanced industrial democracies* (pp. 3–16). Oxford, UK: Oxford University Press.
- Duverger, M. (1954). *Political parties: Their organization and activity in the modern state* (B. North & R. North, Trans.). London: Methuen. (Original work published 1951)
- Duverger, M. (1972). *Party politics and pressure groups: A comparative introduction*. New York: Thomas Y. Crowell.
- Epstein, L. D. (1967). *Political parties in Western democracies*. New York: Praeger.
- Inglehart, R. (1977). *The silent revolution*. Princeton, NJ: Princeton University Press.
- Inglehart, R. (1997). *Modernization and postmodernization: Cultural, economic and political change in 43 nations*. Princeton, NJ: Princeton University Press.
- Katz, R. S., & Mair, P. (1993). The evolution of party organizations in Europe: The three faces of party organization. In W. Crotty (Ed.), *Political parties in a changing age* [Special issue]. *American Review of Politics*, 14, 593–617.
- Katz, R. S., & Mair, P. (1995). Changing models of party organization and party democracy: The emergence of the cartel party. *Party Politics*, 1, 5–28.
- Katz, R. S., & Mair, P. (2002). The ascendancy of the party in public office: Party organizational change in twentieth century democracies. In R. Gunther, J. R. Montero, & J. Linz (Eds.), *Political parties: Old concepts and new challenges* (pp. 113–135). Oxford, UK: Oxford University Press.
- Kirchheimer, O. (1966). The transformation of the West European party systems. In J. LaPalombara & M. Weiner (Eds.), *Political parties and political development* (pp. 177–200). Princeton, NJ: Princeton University Press.
- LaPalombara, J., & Weiner, M. (1966). The origin and development of political parties. In J. LaPalombara & M. Weiner (Eds.), *Political parties and political development* (pp. 3–42). Princeton, NJ: Princeton University Press.
- Lawson, K., & Merkl, P. (1988). Alternative organizations: Environmental, supplementary, communitarian, and antiauthoritarian. In K. Lawson & P. Merkl (Eds.), *When parties fail: Emerging alternative organizations* (pp. 3–12). Princeton, NJ: Princeton University Press.
- Lijphart, A. (1969). Consociational democracy. *World Politics*, 21(2), 207–225.
- Lipset, S. M., & Rokkan, S. (1967). Cleavage structures, party systems, and voter alignments. In S. M. Lipset & S. Rokkan (Eds.), *Party systems and voter alignments* (pp. 91–138). New York: Free Press.
- Mair, P. (1990). Introduction. In P. Mair & G. Smith (Eds.), *Understanding party system change in western Europe* (pp. 1–22). London: Taylor & Francis.
- Riker, W. (1962). *The theory of political coalitions*. New Haven, CT: Yale University Press.
- Sartori, G. (1976). *Parties and party systems: A framework for analysis*. Cambridge, UK: Cambridge University Press.
- Ware, A. (1996). *Political parties and party systems*. Oxford, UK: Oxford University Press.

19

ELECTORAL SYSTEMS IN COMPARATIVE PERSPECTIVE

WILLIAM M. DOWNS

Georgia State University

Elections are central to the functioning of democratic systems, and as such they have been the focus of extensive political science research for centuries. Scholars and practitioners seek to understand the variation in choices of different electoral systems cross-nationally. They also try to isolate the impact of those choices on a range of individual-, institutional-, and system-level outcomes. Those outcomes include the quality and breadth of representation; size and polarization of political party systems; citizen participation and voting behavior; and government, as well as system, stability. Much of the research conducted on electoral systems and elections has evolved from theoretical and empirical work on the United States and other established Western democracies (especially those in Europe), but considerable effort in recent decades has been devoted to understanding elections in transitioning and new democracies globally. Although elections do take place in nondemocratic polities, they usually fail to be free, fair, or competitive and therefore typically fall outside the domain of comparative research. What is clear is that the increasing sophistication of theoretical and statistical tools available to political scientists (along with an expanding universe of cases against which to test expectations) has resulted in important advances in our understanding electoral systems and elections. Because electoral processes and outcomes exert such profound effects on the real world of

politics, such understanding is an example of the crucial connection between theory and practice in political science.

A first level of comparison identifies the different types of elections designed to determine national executive power. Presidents and other chief political executives may be elected through *direct* or *indirect* means. In direct elections of presidents and presidential-type executives, voters cast ballots for one of the eligible candidates. A candidate can win outright in a single round of voting by garnering an absolute majority of the ballots cast; however, when no one candidate captures 50% plus one of the eligible votes, a runoff round is held at a subsequent date, with the top two finishers from the first round squaring off. This model of direct executive election (exemplified by countries such as France, Russia, Poland, and Argentina) has the putative advantage of providing a wide range of candidates from which voters may choose in the first round. If a second round proves necessary, the system then yields a winner supported by an absolute majority of those turning out for the runoff. Indirect election of a country's president or other executive leader, alternatively, entails voters' selecting other persons (*electors*), who will then determine the winner. The United States, for example, uses an indirect mechanism whereby voters choose presidential electors, who then comprise the *Electoral College*, which then votes on who will become the next president. That process

leaves open the possibility that the person chosen to be president by the Electoral College is not the same person who secured the greatest number of votes among the general population. Elsewhere, directly elected parliaments (either an upper house or both houses in joint session) constitute the arena in which presidents are indirectly selected; this occurs in Germany and Italy, for example. When *Westminster style* parliamentary systems (i.e., those modeled after the British House of Commons) use votes by legislators in plenary sessions to approve (or remove) prime ministers as heads of government, they are engaging in indirect elections of political executives. Taking the selection of national leaders out of the direct control of voters represents the skepticism that constitutional architects have for the general population, and it can provide an apparent elite-level check on the sentiments of mass electorates.

The second major dimension along which political scientists compare elections is the method of voting for legislative assemblies. Indeed, examining legislative elections across countries reveals considerable variation in such key dimensions as district magnitude, electoral formulae, ballot structure, and the use of electoral thresholds. *District magnitude* refers to the number of candidates who will be elected to a legislature from any given constituency, and the basic distinction here is between systems that rely on single-member districts and those that employ multimember districts. District magnitude is usually studied in tandem with the system's chosen *electoral formula*, which represents the particular mechanism for translating votes into legislative seats. Such mechanisms are most frequently of the plurality, majoritarian, and proportional varieties. In the single-member district system, a country is divided into discrete electoral districts from which one individual will emerge as the elected representative. This system normally relies on a plurality rule, meaning that the candidate with the most votes wins (regardless of whether that candidate has captured an absolute majority). As such, single-member district systems are often deemed *first past the post* systems and also constitute a winner-take-all approach that provides no electoral prize for coming in second. The United States and the United Kingdom are among the countries where the single-member plurality system has a long-standing history; however, a range of countries elsewhere—including Canada, Ghana, and India—have adopted the same method. Others, most notably France, employ a single-member district system with two rounds of voting. In such cases, individual candidates can win outright in the first round with an absolute majority of votes cast, or they can secure the plurality of votes cast among eligible candidates in the second-round runoff. Single-member district systems are defended by their advocates as those that can enhance clarity of responsibility and democratic accountability by giving citizens in each district one individual to whom credit or blame can be assigned. The clarity and accountability that are supposed to accompany majoritar-

ian governance should, according to this logic, produce more stable and effective polities. Detractors, however, find that aggregating district-level winner-take-all elections into a national whole can produce skewed representation in the legislature. For example, a party that runs a consistent and respectable second place throughout the country but that fails to win any single district would be excluded from taking seats in the legislature. Such a system, then, has the potential to underrepresent small parties in a democracy.

The alternative to single-member, winner-take-all systems of electing representative assemblies is one based on proportional representation (PR) in multimember districts. In PR systems, the goal is to have the percentage of a party's seats in the legislature reflect the percentage vote share captured by that party in the general election. The party securing 25% of the vote would, accordingly, be rewarded with 25% (either exactly or approximately) of the legislative seats. Here *ballot structure*, which shapes how voters cast their votes, becomes critically important. Ballots can be *categorical* or *ordinal*. The categorical ballot structure allows a single either-or choice of one candidate. By contrast, the ordinal ballot structure gives voters the opportunity to vote for more than one candidate. In some ordinal ballots, political parties devise rank ordered lists of candidates to determine which persons ultimately claim those seats. In this *closed party list system*, citizens vote only for a party in a multimember constituency (often the whole country), whereas in an *open party list system*, voters can choose from a published list or select an individual candidate. The closed party list mechanism clearly vests considerable power in the hands of party leadership. Often, PR systems will set a minimum threshold (5% in Germany, for example) that parties must clear in order to win seats. *Electoral thresholds* are an increasingly common way for PR systems to limit the entry of minor (and sometimes extremist) parties into legislatures. Thresholds normally require a minimum percentage of votes or a minimum number of seats in order for a party to gain seats in a legislature. Thresholds vary, with some countries opting to set the bar low (Israel, for example, at 2%) and others raising it to high levels (e.g., Turkey, at 10%). Numerous varieties of proportional representation exist, each with different counting and procedural mechanisms. One such variety is the *single transferable vote* method. By this method, voters rank candidates preferentially, and if a voter's first-choice candidate has already cleared a set threshold and does not need additional support to win, then that vote is transferred to a second choice. This process, exemplified most clearly by Ireland, is designed to avoid "wasting" votes.

Although there is a tendency among political scientists to classify electoral systems in democratic countries into either the majoritarian or proportional camp, the reality is that many hybrid or mixed systems exist in between those two types. The *additional member system*, for example, combines elements of conventional first-past-the-post

systems with some characteristics of party-list proportional systems. In this combination, voters get two votes: The first helps allocate seats to single-member constituencies, and the second goes to a party list. The percentage of second or party-list votes won by a party determines the party's overall number of representatives, and the number of seats won in single-member districts is "topped off" to match that overall percentage. This method finds use in elections to Germany's Bundestag, New Zealand's House of Representatives, and the Scottish and Welsh Assemblies in the United Kingdom. The presumed advantage of this *mixed member system* approach is that proportionality is ensured, and at the same time, a directly accountable representative for each constituency is also identified. It is also said to allow strategic voters to express support for an individual politician while not necessarily endorsing that candidate's political party. Disadvantages are said to include the creation of two (potentially unequal) classes of politicians, with those elected under the second-ballot topping off beholden not to the voters but to party leaders instead.

While elections in democratic settings constitute the overwhelming preponderance of all voting processes studied by political scientists, it is important to note that nondemocratic systems (e.g., authoritarian and semiauthoritarian systems) can also employ electoral mechanisms. Such regimes may organize controlled and uncontested elections as a means of mobilizing mass endorsement of a national leader or a single-party legislature. Doing so can provide symbolic legitimacy for the ruling elite, and it may neutralize popular discontent by creating the false appearance of citizens having a say in the affairs of their country. For example, while the Communist Party monopolizes power and controls political processes in China, direct elections of village-level offices do take place, as do indirect elections for people's congresses above the local level. The one-party Soviet Union held its own brand of uncontested elections, as did Iraq under Saddam Hussein. Brazil under military rule orchestrated compulsory voting in tightly controlled elections, even though the frequency of blank and spoiled ballots often suggested popular rejection of the process. Semicompetitive, hegemonic party systems such as Egypt's hold elections in which there is little *a priori* uncertainty of the outcomes; there is, in such cases, some element of choice and voter expression. Although nondemocratic variants of the electoral process illustrate more about a regime's methods of system control than they do about representation, responsiveness, and accountability, they clearly merit attention.

Theory

Some democratic theorists view elections as a central—if not *the* central—component of liberal democracy.

Indeed, in this view, elections constitute the minimum necessary requirement for democracy. For Schumpeter (1942), democracy is "that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote" (p. 269). Likewise for Huntington (1993), democracy is defined most essentially by the fair and periodic voting procedures that select a country's leaders. Przeworski, Alvarez, Cheibub, and Limongi (2000) also view contested elections—that is, those in which there is *ex ante* uncertainty and *ex post* irreversibility—as the litmus test for democracy. Others, such as Dahl (1971), counter that such a thin, minimalist, or procedural definition of democracy-as-elections fails to account for other necessary conditions, such as the protection of civil liberties and the actual responsiveness of government policies to voter preferences. Whether sufficient or not, elections typically figure as necessary conditions for the existence of democracy.

Theoretical work on elections and comparative electoral systems has largely focused on (a) the relationship between electoral rules and the size and polarization of political party systems, (b) the tendency of certain electoral systems to impact voter turnout and citizen participation, (c) the congruence between electoral verdicts and government policy, and (d) the potential for electoral systems to predispose new or transitioning systems to success or failure. Political scientists developing theory in each of these areas represent some of the main ontological camps in the discipline, such as structuralists, rationalists, and culturalists. As such, attention has been devoted to formal rules, voter preferences and behavior, and the contextual influence on system choice and outcomes.

The causal relationship between electoral rules and the nature of a country's political party system has animated scholarly interest for decades. Perhaps the most famous proposition, tested repeatedly since its early assertion by Duverger in 1954, is that plurality elections using one-ballot single-member districts will favor the creation of two-party systems whereas proportional representation rules with multimember districts will lead to multiparty systems. Duverger went further to posit that a majority vote on two ballots increases the likelihood of a multiparty system as well as the necessity of postelection coalition formation. It is rare indeed that causal relationships in political science theory elevate to lawlike status, but in this case, "Duverger's Law" has achieved considerable staying power. The logic guiding Duverger's assertions depends on what are conventionally deemed *mechanical effects* and *psychological effects*. The mechanical effects highlight the underrepresentation of third (and fourth, and fifth, etc.) parties, which is likely to occur over time in a single-seat legislative district requiring an outright plurality or majority vote. Given the mechanical impediments to minor party success, voters who support minor parties then have psychological incentives not to "waste" their votes and may often cast

ballots against their preferred candidate in a strategic effort to exercise some influence over the most likely winner in the two-party competition. Such claims have spawned much subsequent work, and not a little dissent. Sartori (1968) extended Duverger's assertion of a link between proportionality and party system size, specifying that district magnitude (i.e., the number of seats in a district) is the single best predictor of the effective number of political parties in a district. Riker (1982) challenged Duverger's hypothesis about PR and multipartism by contending that, if true, we should see a recurring increase in the number of parties over time rather than party system stability or modest decreases in the effective number of parties (as most frequently occurs in practice). Debate over the relationship between choice of electoral system and party system size is important, given the propensity to view two-party majoritarian countries as more stable than those with polarized multipartism.

If electoral rules biased in favor of two-party systems are theorized to bring gains in terms of system stability, then those rules favoring proportionality figure prominently in political science theories that attempt to explain citizen engagement, voter turnout, and representativeness of legislatures. According to Lijphart (1994, 1999), majoritarian and plurality electoral systems dilute citizen enthusiasm and voter turnout because so many supporters of minor parties conclude that casting their ballots will have little to no impact on electoral outcomes, government formation, or policy choices. Conversely, proportional systems with low thresholds for representation and large district magnitudes should increase the chances that smaller parties from across the ideological spectrum will be able to secure voice and seats in the legislature. With that greater likelihood of electoral success for minor and even fringe parties, voter efficacy and incentives to cast ballots should be increased. Voter turnout is "an excellent indicator of democratic quality" (Lijphart, 1999, p. 284), and PR systems are theorized to be superior to their majoritarian counterparts in generating democratic gains in this area. As part of an overall inclusive and consensual approach to democratic governance, proportional electoral systems should also improve citizen satisfaction with the political system, *ceteris paribus*.

A third major area of theoretical work on comparative electoral systems has evolved around the presumed correspondence between voting outcomes and public policy. If democracies are to be responsive to the preferences of the public, then periodic voting should work to translate the "will of the electorate" into identifiable policy choices. Scholarship in this area builds on the majoritarian-proportional dichotomy to examine citizen control over—and influence on—government policy making. Powell (2000) explores elections as "instruments of democracy" and distinguishes a proportional vision of "citizen influence" from a majoritarian vision of "control." He contends that "proportional influence designs enjoy a surprising

advantage" (p. 18) over the majoritarian alternative because they encourage broad cross-party bargaining to form a government and to pass legislation. Such bargaining should produce governments that include the *median legislator*, who is, in turn, close to the *median voter*. The median voter is located at the middle of a political system along most issue dimensions, such that one half of the electorate is positioned to the political left and the other half is positioned to the political right. The median legislator is likewise the elected representative located such that half of the other legislators are to the left and the other half are to the right, politically. Electoral systems that produce governments proximate to the median voter should, therefore, be more responsive to policy preferences. Proportional electoral systems should also give greater policy influence to opposition parties, making for a more inclusive process of policy making.

Theories underpinning our understanding of electoral rules and their consequences can have extremely important practical applications. While much effort is devoted to understanding how and why established democracies tinker with their electoral systems to enact reforms or alter a range of political outcomes, even greater attention has been directed in recent decades to the role of elections in facilitating regime change. Indeed, one of the growth areas in political science literature addresses the prospects for successful *electoral engineering*. Given that the last decades of the 20th century witnessed transitions from communism, apartheid, and other forms of autocratic rule, alternative theories about the prospects for implanting democracy through institutional engineering have become increasingly salient. Likewise, nascent postauthoritarian systems in early 21st-century hot spots such as Iraq and Afghanistan have emerged as testing grounds for the discipline's theoretical assertions. Norris (2004) identifies two theoretical traditions—*rational choice institutionalism* and *cultural modernization*—that purport to explain the possibilities for electoral engineering on human behavior. In the rational choice institutionalism approach, political parties adopt discernibly different strategies based on the nature of electoral thresholds and ballot structures. Preference-maximizing citizens likewise should be expected to respond differently to alternative electoral rules. If correct, this logic would predict that rule-based incentives will shape consistent patterns of behavior; therefore, changing those incentives through electoral engineering "should have the capacity to generate important consequences for political representation and for voting behavior" (Norris, 2004, p. 15). By contrast, the cultural modernization approach suggests that deep-rooted cultural habits arising from processes of social modernization place real limits on the potential of formal rules to alter behavior in systematically meaningful ways. This culturalist argument is often employed to explain why wholesale introduction of electoral rules into culturally divided, postconflict settings so

frequently fails to produce short-term transformations of individual behavior.

Applications and Empirical Evidence

Political scientists have endeavored to assemble an abundance of empirical evidence in support of their theoretical claims. Perhaps nowhere has greater effort been extended than in tests of propositions about the linkages among electoral laws, party systems, and coalitional incentives. Countering an alternative hypothesis that underlying societal cleavages are the primary agents determining size and polarization of party systems, a literature has evolved (cf. Cox, 1997; Lijphart, 1994; Rae, 1967; Riker, 1982; Sartori, 1968; Taagepera & Shugart, 1989) to contend that electoral laws have their own independent effects. Duverger's notions about first-past-the-post, single-ballot elections tending to produce two-party majoritarian systems find extensive application in the United States, as well as the United Kingdom. In elections for the U.K. House of Commons and the U.S. Congress, the evidence seems to suggest a compelling link between electoral rules and strong, stable, two-party government. Electoral structures in the United States, for example, help explain the consistent failure of third parties to mount successful campaigns. This winner-take-all system has, though, placed such significant importance on the drawing of district boundaries that the pernicious practice of *gerrymandering*—consciously redrawing the lines to ensure a majority for one party—emerged as part of American politics. Although smaller parties have been able to win parliamentary seats in the United Kingdom, their ultimate representation in the House of Commons is highly disproportionate to their overall support in the electorate, and they have little chance at becoming the party of government or forcing a coalition. To illustrate, the perennial third-party Liberal Democrats won 22.1% of the vote in Britain's 2005 general election but secured only 9.6% of the 646 seats in the House of Commons. Tony Blair's Labour Party, having won only 35.3% of the votes nationwide, nevertheless captured 55.2% of the seats in parliament and 100% of the cabinet positions in government.

Single-member-district plurality systems normally provide rapid certainty after an election about who will govern and who will constitute the opposition. However, systems that introduce even a modicum of proportionality likewise introduce an element of uncertainty into the government formation process. Proportionality (especially when combined with low thresholds in multimember districts) does increase the number of effective parties in the political system. When no single political party secures an outright legislative majority, the postelection period becomes one marked by formal negotiations as well as backroom deals between parties jockeying to join a governing coalition. The case of Belgium is illustrative. There

a proportional representation system with compulsory voting and a 5% threshold for representation in the federal Chamber of Representatives produced enough support to grant parliamentary seats to 11 parties in the June 10, 2007, general election. The largest among those, the Christian Democratic and Flemish Party, claimed only 18% of the 150 seats in parliament and could therefore not form a government by itself. Protracted negotiations commenced after the election, and 196 days later, the best the Belgian parties could do was constitute an interim caretaker government. That interim government lurched along, with further negotiations taking another 79 days before the parties could agree on a full-fledged new government. That government, in turn, failed to finish out the year. Although electoral rules biased in favor of majoritarianism typically yield governments that combine certainty with disproportional representation, those rules favoring multiparty outcomes tend to better reflect the dispersion of political preferences throughout the country but may also add considerable uncertainty to the government formation process.

Evidence also exists on the relationship between electoral systems and the production of such democratic goods as high voter turnout and citizen satisfaction. Where the electoral rules reduce the costs (e.g., time and effort) to citizens of registering and voting, we should find greater turnout. Similarly, where party choices available to voters are more extensive we should expect to see elevated turnout. Finally, voter efficacy—the belief that casting a ballot can actually impact the government formed and the ultimate policy direction taken—should be directly related to turnout at elections. According to Norris (2004), “Institutional rules do indeed matter: voting participation is maximized in elections using PR, with small electoral districts, regular but relatively infrequent national contests, and competitive party systems, and in presidential contests” (pp. 257–258). There is also evidence to support theoretical contentions that the type of electoral system can impact the opportunities for women and minorities seeking to earn a legislative seat or executive office. Among established democracies, the countries that consistently sit atop comparative rankings of the proportion of women winning seats in national parliaments are Sweden, Iceland, Finland, and the Netherlands. Each country employs some form of proportional electoral rules with low thresholds, and in each it is routine for women to constitute more than 40% of national parliamentary representation. Findings such as this, it should be noted, must also take regional political culture and other potentially intervening factors into consideration.

In his study of democratic performance in 36 countries from 1945 to 1996, Lijphart establishes empirically that electoral systems favoring consensus-oriented governance yield gains in citizen satisfaction. When the rules of the electoral process encourage multipartism and coalition building, the policy preferences of the median voter have a

greater chance to be represented in the government of the day. Lijphart's data show that the distance between governments and median voters is highest in majoritarian systems (with the United Kingdom representing the high end of the scale) and lowest in more proportional systems (with Ireland and its single transferable vote system producing the narrowest gap). Because in PR systems electoral "losers" often have a chance to join postelection coalitions—and due to the frequent proportional representation of opposition parties on legislative committees—Lijphart is able to find a statistically significant difference between citizen satisfaction in countries with alternative electoral systems. Lijphart's study corroborates earlier work by Klingemann (1999), who found that Danes and Norwegians—each with highly proportional systems—scored the highest levels of democratic satisfaction.

Given the volume of empirical applications of existing political science theoretical work on elections, it is not surprising that there is a foundation of cases demonstrating how changes in electoral rules actually impact voter behavior and system characteristics. Indeed, the lessons of major 20th-century electoral reforms in three countries—France, Japan, and New Zealand—are instructive. The French case illustrates how constitutional architects can try to contain what are perceived to be the excesses of proportional representation. Those designing the 1958 Fifth Republic sought to use electoral rules to avoid reproducing the fleeting and weak multiparty coalition governments that had plagued the Fourth Republic from 1946 to 1958 and brought the system to the brink of collapse. The new two-round, single-member district system established in 1958 encouraged broad political party competition in a first round and awarded National Assembly seats to all candidates winning an outright majority. Absent a majority, all candidates receiving at least 12.5% of first-round votes could then contest in the runoff election, in which a plurality would suffice for victory. In practice, this runoff mechanism encourages the weakest candidates to voluntarily stand down in favor of a better positioned candidate closest to them on the left-right ideological spectrum and to have their supporters cast their second-round ballots for that person. This system has effectively preserved France's multiparty system while simultaneously creating a stable two-bloc system of parties on the moderate left and right. The runoff system often means that parties with meaningful support nationwide may still fail to secure national legislative seats, as has been the case with the far-right National Front party. Indeed, when the French tinkered with their electoral laws in the 1980s, it became apparent how decisive the rules can be for representation. In 1986, the Socialist government of President François Mitterrand opted to change from the two-round system to a single-round proportional one in hopes of dividing the right wing opposition parties. As a result, the National Front's 9.6% of the vote earned it 35 of the 577 national legislative seats. When party strategy changed and France reverted to

the two-round system for the 1988 parliamentary election, the National Front's 9.7% of the first-round vote translated into only one seat!

In Japan, major reforms occurred in 1994, when the old system of single nontransferable votes (allowing one choice per voter in elections for three to five district representatives) was scrapped and replaced by a mixed-member system. The new Japanese system for electing the House of Representatives combines first-past-the-post single-member districts (for 300 seats) with PR party-list seats (200) in an "attempt to craft a competitive two-party, issue-oriented politics and a cleaner, more efficient government" (Norris, 2004, p. 5). Whereas Japanese politics prior to the reform consisted mainly of one dominant party (Liberal Democrats) regularly overwhelming a handful of opposition, the new hybrid of majoritarian and proportional approaches (most analogous to the system in Russia) aims to create a polity with alternating parties in power. In New Zealand, at roughly the same time, reforms to replace the long-standing first-past-the-post system came to fruition. There, a mixed-member proportional system now allows 70 of the 120 national parliamentary seats to be elected directly in single-member districts, with the remainder coming from party lists in a style similar to Germany's. The addition of proportionality to New Zealand's electoral system—endorsed by a majority of citizens in a binding 1993 referendum—has had a quick and dramatic impact. Whereas the average number of political parties gaining seats in New Zealand's national parliament was just two during the 1946-to-1993 period, in the five elections since introducing the mixed-member system, an average of seven parties has secured representation. Electoral engineering, at least in this case, seems to have achieved the end envisioned for it.

Perhaps nowhere is political science research into comparative electoral systems more salient than in countries attempting to transition away from authoritarianism. The cross-national lessons available to architects of new systems are always imperfect, as transporting a model from one country to another without sensitivity to local conditions and histories is a formula for failure. However, such comparative learning does take place, and most new electoral systems today are adaptations and amalgams of those found elsewhere. When elections were held in December 2005 to constitute a post-Saddam Iraqi Council of Representatives, a proportional party-list system determined 230 of the total 275 seats in 18 multimember districts (*governorates*). An additional 45 compensatory seats were then allocated to political entities that did not win any seats outright in the governorates but that did clear a minimum national threshold. Also worthy of note is that Iraq's electoral law requires at least 25% of the members of the parliament to be women. In Afghanistan, post-Taliban elections have struggled to secure domestic and international legitimacy. The 2005 elections for Afghanistan's lower house of parliament employed the single nontransferable

vote method in 34 multimember constituencies. Candidates, however, ran independently because parties and lists were not recognized by the governing law. As in Iraq, the Afghan system reserved a number of seats (at least 68 of the total 249) for women. At the executive level, the Afghan president is elected by absolute majority in a two-round system similar to that employed in France.

Policy Implications

The choice of election system can potentially impact the quality and kind of policy pursued by an incumbent government. If elections are the essential ingredient in representative democracy, then presumably there should be some apparent connection between the will of the people as expressed through elections and the policies they receive from the subsequently invested government. If citizens are engaging in issue *voting*, as some research has consistently found, then it is important to gauge whether the governments they get are actually responsive to those issues. If our fundamental expectations about democracy require a close connection between elections and policy outcomes, then the reality may sometimes disappoint (Ginsberg & Stone, 1996). As Downs (1957) contended, political parties adopt policies in order to win elections rather than win elections in order to adopt policies. The achievement of public policy goals may actually be instrumental to the more power-seeking ambitions of parties and politicians.

Political scientists therefore examine the ways in which different electoral systems hold officials accountable for their fidelity to campaign promises once in office. Indeed, elections provide a kind of *ex post* accountability for policy pledges. The more a particular model of election creates the perception among elected officials that those they claim to represent will oust them for poor past performance, the stronger the democratic accountability linkage is said to be. Although several scholars (e.g., Lijphart, 1999; Powell, 2000) demonstrate that citizen satisfaction and the correspondence between median voters and the policy positions represented in a legislature are enhanced by consensual, proportional representation, there is also reason to find that majoritarian systems provide the kind of clarity that voters need to hold leaders accountable for policy choices. In a majoritarian, winner-take-all system, if a party campaigns on the basis of very clear policy pledges, wins the election, and then proceeds to depart dramatically from its public promises (what political scientists call engaging in *moral hazard*), the voters should be able to easily identify this lack of fidelity and then “throw the rascals out” at the next electoral opportunity. By contrast, in electoral systems characterized by proportional representation, the likelihood of multiparty coalition governments forming after protracted negotiations is great. In such cases, the

translation of electoral verdicts into governmental policy becomes significantly more indirect. Moreover, the distribution of policy portfolios across multiple parties blurs the lines of accountability and increases the difficulties for voters who wish to reward or punish the incumbents. For example, the citizen asked to evaluate with one vote the performance of a three-party coalition government may find it hard to express support for that government’s fiscal policy (headed by a Conservative Party finance minister) while rejecting its policies on education (headed by a Christian Democratic Party education minister), as well as those on immigration (headed by a Nationalist Party interior minister).

If elections are central to the functioning of democratic political systems, then another set of policy implications can be found in the promotion of democratization through elections and electoral reform. The foreign policies of many established democracies, as well as those of intergovernmental organizations and donor agencies, are intimately tied to this kind of promotion. The conduct of free and fair elections is frequently the litmus test for legitimacy in the eyes of the democratic international community, and everything from diplomatic recognition to commercial relations can hinge on the successful holding of competitive elections. As such, governmental entities such as the United Nations or the European Union will regularly send election monitoring missions to observe voter registration and the casting of ballots to gauge openness, extent of fraud, and incidents of intimidation. Nongovernmental organizations, such as the Carter Center, have also played this monitoring role in Africa, Latin America, and Asia. International financial institutions, such as the World Bank, also incorporate elections into decisions about granting development assistance funds to countries in need. This process of *political conditionality* is the stipulation of the conduct of democratic elections as a necessary occurrence prior to the allocation of foreign aid. Such conditionality has been part of the Structural Adjustment Programs implemented by the World Bank and the International Monetary Fund in exchange for lower interest loans to developing countries. Critics of these policies contend that tethering development assistance to political reforms is tantamount to threats that will lead to the rapid importation of electoral mechanisms that ultimately fail to take root.

Future Directions

As democracy expands (and sometimes contracts) across the globe, research on elections likewise adapts. Political scientists continue to focus on formal rules and designs; on individual-level attitudinal and behavioral responses to those formal mechanisms; and on the connections among elections, party systems, and policy outcomes. One of the particular growth areas for future research in the area will be that addressing referenda and other forms of direct

democracy. Referenda can take many different forms, with some being ad hoc and others constituting routine and regular procedures. They are advocated on the logic that circumventing the normal representative institutions in favor of direct votes by the entire electorate will encourage more citizens to become better informed and more involved in the democratic decision-making process. Singling out a policy choice for a decision by the people should, moreover, grant clarity to the direction desired by citizens; in contrast, the normal process of bundling multiple policy choices within legislative bills makes it hard to achieve such clarity. Finally, it is presumed that decisions arrived at through referendum elections will enjoy much greater legitimacy than those achieved through competition or cooperation of the political elite. With greater legitimacy, we should expect, in turn, the greater likelihood that those policies are successfully implemented. Referenda, plebiscites, and citizens' initiatives take place in many countries: Switzerland uses the people's initiative with relative frequency (more than 400 national referenda since 1945); French and Dutch voters were asked in their respective 2005 national referenda whether they supported a proposed European Union Constitution; and voters in East Timor chose to part from Indonesia in a 1999 referendum. Although the United States does not hold national referenda, some states and many localities do hold frequent initiatives and ballot propositions (for example, California's 1978 Proposition 13 on property taxes and its 2008 Proposition 8 on same-sex marriage).

Future research should build on the increasing relevance of referenda and initiatives to explore the impact of question wording and ballot structure on voting outcomes. Assumptions about referenda being decided by informed and engaged citizens need to be tested thoroughly, across time and across countries. If empirical support does not emerge to substantiate that there are informational gains associated with referenda to a greater extent than in regular elections, then some normative red flags need to be raised regarding the utility of this form of democratic decision making. Fodder for further research comes with the varying turnout requirements that countries impose for the verdicts of direct democracy votes to be implemented: Why do different countries place the turnout threshold at different levels, and do high thresholds unfairly violate majorities that fail to reach them? Do voters behave differently when referendum elections are merely advisory rather than binding on the government? Most important, is there any empirical reason to expect that direct democracy elections are supplanting conventional political elections in any meaningful way? These trends and prospects clearly deserve the attention of 21st-century political scientists.

Conclusion

The rules governing and guiding voting are central to the study of contemporary politics. Decades of comparing

electoral systems have produced important findings about the impact of alternative majoritarian and proportional systems and the many hybrid models in between. Duverger's early assertions about the influence of ballot and district type on the size and character of political party systems have risen to the status of "law" in the discipline and spawned subsequent and more sophisticated theorizing. A considerable body of evidence now exists to help explain how the choice of electoral system can influence the quality of a country's democracy. Introducing or reforming electoral rules can alter citizen participation and satisfaction, can enhance or diminish the congruence between voter preferences and public policy outputs, and can have profound consequences for system stability. Electoral engineering as such is one of the clearest issue areas in which political science research speaks directly to decision makers. Indeed, given that for centuries revolutions have been fought and blood spilt for the right to live in a democracy, it is imperative to understand how elections can support or undermine transitions from authoritarian rule. Some of the most tenuous polities around the globe struggle with legitimacy and leadership transitions, and designing an appropriate electoral system shapes—if not determines—those countries' futures. Theories and data assembled for study of the United States, western Europe, and other cases of consolidated democracy offer much to the electoral engineer, policymaker, and student observer; they cannot, however, be casually transported across the globe to nascent democracies without due consideration of the opportunities and constraints defined by a country's individual context.

References and Further Readings

- Amorim Neto, O., & Cox, G. (1997). Electoral institutions, cleavage structures, and the number of parties. *American Journal of Political Science*, 41(1), 149–174.
- Amy, D. (2000). *Behind the ballot box: A citizen's guide to electoral systems*. Westport, CT: Praeger.
- Anderson, C., & Guillory, C. (1997). Political institutions and satisfaction with democracy. *American Political Science Review*, 91(1), 66–81.
- Bogdanor, V. (1984). *What is proportional representation?* Oxford, UK: Martin Robertson.
- Bogdanor, V., & Butler, D. (Eds.). (1983). *Democracy and elections*. Cambridge, UK: Cambridge University Press.
- Cox, G. (1997). *Making votes count: Strategic coordination in the world's electoral systems*. Cambridge, UK: Cambridge University Press.
- Crewe, I., & Denver, D. (1985). *Electoral change in Western democracies: Patterns and sources of electoral volatility*. New York: St. Martin's Press.
- Dahl, R. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.
- Dalton, R., Flanagan, S., & Beck, P. (Eds.). (1984). *Electoral change in advanced industrial democracies: Realignment or dealignment?* Princeton, NJ: Princeton University Press.

- Downs, A. (1957). *An economic theory of democracy*. New York: Harper & Row.
- Duverger, M. (1954). *Political parties: Their organization and activity in the modern state*. London: Methuen.
- Farrell, D. (1997). *Comparing electoral systems*. London: Prentice Hall/Harvester Wheatsheaf.
- Farrell, D. (2001). *Electoral systems: A comparative introduction*. New York: Palgrave.
- Ginsberg, B., & Stone, A. (Eds.). (1996). *Do elections matter?* Armonk, NY: M. E. Sharpe.
- Grofman, B., & Lijphart, A. (Eds.). (1986). *Electoral laws and their political consequences*. New York: Agathon Press.
- Huntington, S. (1993). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Katz, R. (1980). *A theory of parties and electoral systems*. Baltimore: Johns Hopkins University Press.
- Klingemann, H. D. (1999). Mapping political support in the 1990s: A global analysis. In P. Norris (Ed.), *Critical citizens: Global support for democratic government* (pp. 31–56). Oxford, UK: Oxford University Press.
- Lijphart, A. (1994). *Electoral systems and party systems: A study of twenty seven democracies, 1945–1990*. New York: Oxford University Press.
- Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty six countries*. New Haven, CT: Yale University Press.
- Lijphart, A., & Grofman, B. (Eds.). (1984). *Choosing an electoral system: Issues and alternatives*. New York: Praeger.
- Norris, P. (2004). *Electoral engineering: Voting rules and political behavior*. Cambridge, UK: Cambridge University Press.
- Powell, G. (2000). *Elections as instruments of democracy: Majoritarian and proportional visions*. New Haven, CT: Yale University Press.
- Przeworski, A., Alvarez, M., Cheibub, J., & Limongi, F. (2000). *Democracy and development*. Cambridge, UK: Cambridge University Press.
- Rae, D. (1967). *The political consequences of electoral laws*. New Haven, CT: Yale University Press.
- Reeve, A., & Ware, A. (1992). *Electoral systems: A comparative and theoretical introduction*. New York: Routledge.
- Riker, W. (1982). The two party system and Duverger's law: An essay on the history of political science. *American Political Science Review*, 76(4), 753–766.
- Rule, W., & Zimmerman, J. (1994). *Electoral systems in comparative perspective: Their impact on women and minorities*. Westport, CT: Greenwood Press.
- Sartori, G. (1968). Political development and political engineering. In J. Montgomery & A. Hirschman (Eds.), *Public policy* (pp. 261–298). Cambridge, MA: Harvard University Press.
- Schumpeter, J. (1942). *Capitalism, socialism, and democracy*. New York: Harper & Row.
- Shugart, M., & Wattenberg, M. (Eds.). (2001). *Mixed member electoral systems: The best of both worlds?* Oxford, UK: Oxford University Press.
- Taagepera, R., & Shugart, M. (1989). *Seats and votes: The effects and determinants of electoral systems*. New Haven, CT: Yale University Press.
- Wall, A., Ellis, A., Ayoub, A., Dundas, C. W., Rukambe, J., & Staino, S. (2006). *Electoral management design: The international IDEA handbook*. Stockholm, Sweden: International Institute for Democracy and Electoral Assistance.

20

COMPARATIVE FEDERALISM, CONFEDERALISM, UNITARY SYSTEMS

WILLIAM M. DOWNS

Georgia State University

One of the classic challenges of political organization is the territorial or spatial division of power. Just as constitutional engineers, politicians, and philosophers have struggled over the concentration or separation of horizontal (e.g., executive, legislative, and judicial) powers, so too have battles been fought over the distribution of vertical (e.g., central, regional, and local) authority. Although the late 20th and early 21st centuries have featured significant movement in favor of boundary broadening and the interdependence of outward-looking states, the salience of interactions among countries' internal units remains high. Indeed, understanding the contemporary nation-state demands that political scientists make sense of the centrifugal pressures of decentralization that coexist alongside centralizing trends of integration and unification. Scholars have generally classified approaches to the geographic dispersion of governmental authority by grouping states into unitary, federal, and confederal types. Research on the different system types has evolved from early work on nation building and pacification of regional tensions to contemporary efforts to explain differences in the quality of representation, in the durability of alternative models, and in the adaptability of structural designs to the changing demands of global interdependence.

Unitary systems are those in which sovereignty, decision-making authority, and revenue-raising powers are

clearly and solely vested in a single central government. Subnational units may exist in unitary states, but they enjoy only those powers specifically delegated or "devolved" to them by the central government, and those powers can be revoked at the center's discretion. The majority of the world's states today exemplify characteristics matching this definition of the unitary system. Among those are France (the "one and indivisible" state long held up as a quintessential example of the unitary model), Japan, and China. The United Kingdom is frequently cited as a striking example of a unitary state that has devolved limited powers to a series of regional assemblies (Scotland, Wales, Northern Ireland) without transforming into a federation. Among the putative advantages of the unitary state system are clarity in the lines of accountability, coordinated control by a primary legislature and executive, and the ability to ensure equality in treatment of all parts of the country through uniform application of common laws. Chief among the supposed limitations of the unitary state design are difficulties in accommodating local differences and the potential for excessive concentration of power in the capital.

Federal systems are said to combine self-rule with shared rule, characterized by a strong central government coexisting with subnational units that enjoy their own spheres of jurisdiction and resource bases. Sovereignty is shared across levels of government, and the formal distribution of powers

is defined constitutionally (rather than at the center's discretion). Powers can be taken from subnational units only by amending the national constitution. While no single, pure model of the federation exists, several characteristics are normally central to classifying states in this category. Among these characteristics are the existence of a bicameral legislature (with an upper house designed to represent interests of the constituent subnational units), the presence of a court to adjudicate disagreements between the federal and subnational authorities, and elected executives at the federal and regional levels. The federal (i.e., central) government retains almost exclusive authority to act on behalf of the nation-state in its dealings with foreign entities. Although federations constitute a distinct minority within the universe of contemporary states, they tend to emerge in large or populous countries (e.g., the United States, Canada, Germany, India, Brazil, Australia, Austria) and therefore constitute important influences on a substantial portion (approximately 40%) of the world's citizens. Frequently praised for balancing competing loyalties and facilitating good governance in heterogeneous societies, the federal model is likewise also charged with being excessively redundant and slow due to multiple decision makers ("veto players") and overlapping jurisdictions. Federations are also critiqued for allowing regional and provincial identities to persist as challenges to national unity and for enabling subnational units to pass laws different from those of the federal government.

Confederal systems likewise feature a central government coexisting alongside subnational units, but in this model, the provincial, regional, or state governments are significantly stronger than the national authority. Indeed, the central government relies heavily on the association of subcentral units for resources and the authority to act. Whereas membership of constituent territorial units in a federation is mandatory, in a confederal arrangement the subcentral entities participate voluntarily and can much more easily exit the partnership. Typically, major country-wide decisions rely on the unanimous consent of the regional units. The confederation effectively exists as a community or alliance of individual political entities in which primary emphasis is placed on protecting the liberties and identities of the constituent units. In practice, the central government has a limited policy purpose (primarily defense and foreign policy). The rarest of the state forms, confederation is most frequently associated with the United States under the Articles of Confederation (1781–1789), the Confederate States of America (1861–1865), Switzerland (1291–1847), the Commonwealth of Independent States (former USSR), and the emerging European Union (EU). While generally successful in dealing with a range of common but limited governmental problems, confederations can struggle when weak central authorities cannot enforce national laws, generate sufficient independent revenues, or adjudicate intergovernmental disputes.

Theory

An impressive body of political science literature has evolved regarding the different forms that geographic distribution of power can take. The contemporary literature rests on a foundation of centuries' worth of writings by political philosophers and constitutional architects who focused squarely on both the normative and practical questions of how to build a desired polity. Among those who emerged along this lengthy early chronology to deliberate the proper balance of centralization and decentralization were Althusius, Grotius, Montesquieu, Mill, Hobbes, and Madison. Althusius's *Politica*, written in 1603, is, for example, considered by many to have been the earliest coherent formulation of bicameralism as a method of consensual decision making by communities within a union. Hobbes has often been held up as a defender of the model of centralism (manifested most clearly by the United Kingdom's Westminster system), and Madison is one of the most frequently cited defenders of decentralization as a mechanism for the prevention of tyranny (exemplified by the American federal experience). Their ideas provided inspiration for founders and reformers in multiple country contexts. Whether by imposed force, organic evolution, or explicit choice, models of territorial politics developed with varying degrees of effectiveness and durability. Much of the subsequent theoretical and empirical work on centralism and decentralism by political scientists has focused on the American case, but recent decades have witnessed good comparative work by scholars seeking to identify generalizable patterns from the systematic testing of theory-driven hypotheses. The evolution of theories on unitary, federal, and confederal systems mirrors general trends in the discipline, with formal legalism being supplanted by behavioralist and choice-based approaches.

The early baseline for much of the modern political science work in this area is provided by Riker (1964), Friedrich (1968), and Duchacek (1970). Riker's pioneering book, *Federalism: Origin, Operation, Significance*, was one of the first efforts to move the study of territorial design away from single-country monographs about formal administrative design to a more hypothesis-oriented study of the incentives that motivate politicians to choose one model over another. His theoretical concern was to proffer explanations for two phenomena: (1) adoption of a particular design and (2) maintenance and survival of chosen government forms. In particular, Riker asked, Why adopt a federal model? and Why keep it? Calling a comparative study of federalism "far too pretentious a project for one man" (p. xii), Riker focused his study on the United States while explicitly generating hypotheses that could be tested elsewhere. Conscious of his role as a neutral social scientist, Riker also deviated from conventional practice and decided to minimize normative questions of morality and efficiency in favor of theory-building empiricism. The

crux of his theoretical contribution lies in identification of two factors—the *expansion condition* and the *military condition*—that are necessary to any understanding of comparative federalism. According to Riker, the adoption of a federal model over a unitary form can be explained in part by politicians' concluding that giving concessions to the rulers of constituent units is the only way to secure a desired expansion of power without the use of force. This logic of consenting to shared rule in order to secure expansion constitutes one of Riker's necessary conditions for the *federal bargain*. The second necessary condition is the presence of "some external military-diplomatic threat or opportunity" (p. 12) that would lead politicians to give up some independence in favor of a federal union. Beyond theorizing the factors that lead to adoption, Riker argued that the structure of a country's political party system is what "encourages or discourages the maintenance of the federal bargain" (p. 51). In doing so, he rejected traditional formal-legal explanations based on the particular division of administrative responsibilities. Indeed, for Riker, the theoretical linchpin for understanding the survival of federal forms is the existence of decentralized party systems that can protect the integrity of subnational units in the federal system and thereby help preserve the federal bargain. Riker also placed emphasis on simultaneous citizen loyalty to both federal and regional levels of government, but he contended that loyalty to "dissident provincial patriotisms" is not essential. Having ventured this framework for understanding federalism, Riker nevertheless concluded that notions of federalism as a guarantor of freedom are false.

Friedrich's 1968 book, *Trends of Federalism in Theory and Practice*, marked an additional step in the early development of political science approaches. Even more explicitly than Riker, Friedrich argued that federalism must be seen as a dynamic process embedded in historical tensions rather than as a static institutional design governed by unalterable rules. According to this logic, federalist beliefs and behavior are every bit as important as formal constitutional design and juristic preoccupations with questions of sovereignty, the distribution of competences, and institutional structures—if not more so. Friedrich posited and confirmed the hypothesis that federalism does not work when it is imposed, and he placed the greatest analytical emphasis on an underlying "federal spirit" necessary to maintain the system. The particular form that decentralism through federalism takes—as well as its likelihood of success—"depends on the degree of differentiation in the community, the urgency of the common task, the strength of the interests and beliefs in their particular mix of time and place" (p. 174). Although stopping far short of seeing federalism as a panacea, Friedrich deems it "a useful instrumentality for good government" (p. 184).

Duchacek's *Comparative Federalism: The Territorial Dimension of Politics* (1970) raised further theoretical challenges, mapping out a systematic case against classifying political systems in "rigid unitary, federal, or

confederal categories" (p. ix). Indeed, Duchacek doubted the analytical utility of distinguishing between federal and unitary states and worried that *federalism* could now be described in the following terms:

one of those good echo words that evoke a positive response but that may mean all things to all men, like democracy, socialism, progress, constitution, justice, or peace. We see the term applied to almost any successful combination of unity with diversity, to almost any form of pluralism and cooperation within and among nations. (p. 191)

Admitting that no accepted theory of federalism exists, Duchacek questioned the merits of using the United States as a baseline for all comparisons (but then offered no superior alternative) and disagreed with Riker on viewing federalism as a mutually agreed on bargain:

Rejection of a unitary system and the adoption of some variant of a federal formula is rarely a matter of free choice on the part of a handful of constitutional lawyers gathered in an ivory tower seminar on comparative constitutional law. There is often simply no practical alternative to the adoption of federalism. (p. 206)

Most important, Duchacek's (1970) lasting theoretical legacy lies in creating the "ten yardsticks of federalism" (pp. 112–155) that presume to capture the reality and indestructibility of federal unions. These include measures of central control over national defense and international diplomacy, immunity against dissolution by secession, ability of the central government to exercise its authority independent of approval and resources of constituent units, control over amendments to the constitution, protections for the identities of component units, equal representation of unequal units in a bicameral legislature, independent sets of courts at federal and subnational levels, an independent court to adjudicate disagreements between central and subcentral units, retention by lower levels of those powers not expressly granted by the constitution to the central authority, and clarity in the territorial division of authority.

Together, the early works on federal alternatives to unitary states spawned subsequent research that is increasingly sophisticated in theoretical approach and more broadly comparative in data collection. Of that literature it can be said first, however, that conceptual debate still dominates. Although political scientists remain clear on what constitutes a centralized unitary state, decentralizing trends in formally unitary states have "shifted scores of formally unitary countries into the gray zone between federalism and unitarism" (Eaton, 2008, p. 666). Federalism itself is subject to distinction and disaggregation, as illustrated by King (1982), who refers to *federalism* as ideology or philosophy and to *federation* as institutional fact, clearly suggesting that it is possible to have federalism without federation. Forsyth (1981) views confederation as

a union of states within a single polity, to be distinguished from a federation as a union of individuals within a single polity (the latter representing greater national unity). Beyond conceptual wrangling, however, there are continuing disputes over the relative virtues of different forms of decentralization. Gerring, Thacker, and Moreno (2005) construct a theory of *centripetalism* distinct from centralism that counters a growing scholarly consensus emerging around the merits of decentralization. The diffusion of power among multiple independent decision makers, with power flowing outward in a centrifugal fashion toward the periphery, is “the reigning paradigm of good governance in academic and policymaking circles at the turn of the twenty-first century” (p. 569). The decentralist model, they claim, may, however, actually undermine the quality of governance in a polity by introducing excessive fragmentation, mixed messages, and chaos. A centripetal design, by contrast, can strengthen national political parties, institutionalize corporatist-style interest representation, encourage collegial decision making, and provide authoritative public administration. Such an approach stands in contrast to those who, like Lijphart (1999), see the centripetalism of unitary states as part of a majoritarian model that is excessively exclusive, competitive, centralized, and adversarial.

Applications and Empirical Evidence

Much empirical work has been assembled in attempts to explain why unitary, confederal, and federal forms emerge, endure, and sometimes collapse. As argued by Tarr (Kincaid & Tarr, 2005), those questions and answers find specific form in a range of empirical markers that political scientists can and should compare: adherence by governments and politicians to the regime’s constitutional provisions, the political and social (in)stability produced by the choice of territorial governance model, popular support and satisfaction, democratic performance, provision and guarantees for individual and/or communal rights, and economic performance.

Contemporary cases demonstrate that the durability of unitary forms—still the most frequent arrangement for territorial governance—requires adaptation. Indeed, many countries long considered exemplars of centralized governmental structure have in recent decades engaged in decentralization reforms while stopping well short of committing to a federal bargain. Adaptation through decentralization has been a function of multiple influences. Economic contraction beginning in the 1970s, for one, spawned a number of reforms by central governments aimed at “offloading” burdens to regional and provincial units. Elsewhere decentralization became an instrumental part of oppositional politics—out parties struggling to win office found that pledges to dismantle unitary states could produce gains at the ballot box, and once in office, some of those newly

incumbent parties delivered on campaign promises. Decentralization has also been seen more broadly as a natural progression in the democratization of states that no longer required the tight control necessary at their founding. Other cases reveal decentralization to be driven by the desire to pacify persistent—and sometimes volatile—regional nationalisms. Further, the opportunities afforded by new forms of supranational governance (the EU, for example) have emboldened and empowered subnational entities that see the sovereignty of the traditional nation-state as compromised and increasingly obsolete.

The range of motivations driving the “devolution revolution” (Hueglin & Fenna, 2006) is illustrated clearly by three European cases: the United Kingdom, France, and Italy. In the United Kingdom, the catalyst for devolution was an out party that had lost four consecutive general elections but which seized strategically on decentralization as a popular campaign pledge in 1997 and then followed through on its promises once in power. For centuries, the struggle in what became the United Kingdom of Great Britain and Northern Ireland was to pacify rival national identities in Scotland, Wales, and Ireland and to link them tightly under one crown and one authoritative government in London. Administrative unity, however, never wholly eliminated regional identities and interests. Discovery of North Sea oil in the 1970s—with its potential for revenues—fueled a resurgent Scottish nationalism. Frustration over a disappearing language and cultural distinctiveness powered (albeit to a lesser extent than in Scotland) Welsh demands for autonomy. In Northern Ireland, outright and protracted conflict between Protestant unionists and Catholic republicans spawned *the Troubles* and defied easy resolution. Structural reform through devolution came to each of these constituent units of the United Kingdom when Britain’s Labour Party—struggling in opposition for 18 years—promised in a successful 1997 election campaign to grant powers to directly elected assemblies in Edinburgh, Cardiff, and Belfast. With devolution as a centerpiece of Prime Minister Tony Blair’s new government, Northern Ireland’s elected Assembly convened in 1998 (only to be suspended by London on multiple occasions from 2002 to 2007), and new Scottish and Welsh parliaments were elected in 1999. With devolution initiated, there has been no subsequent evidence of the United Kingdom’s moving toward federation despite the laments of its critics. Indeed, support for preserving the sovereignty of this multination-state has to date overshadowed any emergence of the federal spirit on which Friedrich’s early work placed so much emphasis.

Devolution in the unitary French state illustrates the (especially) left wing ideological belief in democratization through decentralization. French Socialists, led by President François Mitterrand, used their landslide 1981 election to dismantle Parisian centralism and create 26 directly elected regional councils (each with an indirectly

elected president). The decentralization laws of 1982 and 1983 were designed by the Socialist-led left (and over the objections of the Gaullist right) to “lift the rigid Napoleonic iron corset from the statist French bureaucratic government administration” (Tiersky, 2002, p. 148). Adding a genuine bottom-up dimension demonstrates the “realization that even the most famously unitary state of Europe would benefit from some form of shared or divided governance” (Hueglin & Fenna, 2006, p. 23). In Italy, social pressures from below were certainly present in the 1970 reforms that created new administrative regions and devolved powers to them; however, the desire by central authorities to contain expenditures and respond to fiscal pressures was more determinative. Offloading and controlled burden sharing in the Italian system have resulted in marked variations in institutional success throughout the country (e.g., greater success in the north, less in the south). Research into the role of *social capital* and civic traditions in explaining the variable success of Italy’s devolution forms the basis of Putnam’s *Making Democracy Work* (1994), one of the most celebrated political science books in recent decades.

Although some states have demonstrated flexibility and adaptation while retaining their essential character as unitary polities, others have transformed into either full-fledged or nascent federations. Belgium, for example, is illustrative of Duchacek’s (1970) contention that sometimes there is simply no practical alternative to the adoption of federalism. A classic unitary state in the Napoleonic conception since its creation in 1830, Belgium grew into an awkward amalgam of linguistic communities that simmered to a crisis in the late 1960s. A French-speaking south (Wallonia), a Dutch-speaking north (Flanders), and a predominantly francophone capital, Brussels, embedded within the Flemish north meant that cultural and economic disagreements could not be easily resolved through partition or separation. Constitutional reforms in 1970 and 1980 laid the groundwork for the 1993 addition of a new Article 1 of the Belgian Constitution, formally proclaiming Belgium to be a federal country with elected parliaments in Flanders, Wallonia, and Brussels, as well as a reformed national parliament. Deschouwer (2005) described Belgium’s metamorphosis from a unitary state to a federal one as follows:

the result not of a deliberate choice but of incremental conflict management. . . . Federalism just happens to be the system of government that emerged, to some extent as the unwanted consequence of the search for a way to keep two increasingly divergent parts of the country together. (p. 51)

Another “new, although reluctant, member of the family of federal polities” (Steytler, 2005, p. 312) is South Africa. The negotiated compromise reminiscent of Riker’s federal bargain that reshaped the South African state is found in the arrangements for cooperative government through decentralization in the 1996 postapartheid constitution.

Architects of South Africa’s new nonracial democracy sought to simultaneously build a sufficiently strong central government to transform a system tormented by decades of White minority rule while limiting the potential abuses of majority rule through decentralization to new provincial authorities. Granted few exclusive powers, South Africa’s nine provinces enjoy concurrent powers in such important areas as agriculture, environment, culture, health services, housing, transportation, and education. As in established federations such as Canada and India, any remaining residual powers rest with the national government. Replete with federal elements without formally declaring itself a federation, South Africa demonstrates a relatively strong commitment to cooperative decision making among communities. Such is rare in sub-Saharan Africa, with the potential exception of Nigeria.

Empirical applications of confederal principles are more difficult to locate in the 21st century. However, if most political scientists agree on a major political system meeting the standards of strong local units allied in cooperation with a weaker central authority, then that system is the contemporary EU. Majone (2006) deems the EU to be “a failed federation but a successful postmodern confederation” (p. 122). Its 27 independent member states are by treaties linked through common institutions and rules—thereby relinquishing portions of their sovereignty—in order to secure the common goals of economic integration and democratic peace. The member states retain most of their long-standing jurisdiction over domestic and foreign policy, and they are the central actors in generating revenue for the Union. The Council of Ministers, bringing together representatives of the member states’ incumbent governments to determine policies for the EU, is an explicitly intergovernmental institution. Moreover, major changes to the EU (including the addition of new members) cannot take place without the unanimous agreement of all member states. One particularly poignant reminder that the EU is not a fully integrated federal polity is that both Britain and France—not Brussels—retain complete control over their nuclear arsenals. Likewise, neither Britain nor France has relinquished its permanent seat on the United Nations Security Council in favor of the EU. In terms of Duchacek’s other yardstick items for measuring federation, the EU also falls short for failing to have the ability to exercise its authority independent of approval and resources of constituent units, for the absence of an elected bicameral legislature, and for being unable to codify its governing framework in a genuine constitution. In some very real ways, then, the EU appears as a compact between independent units rather than as a fully formed federation.

To be sure, some observers contend that the EU has developed well beyond its confederal origins. For Verney (2002), “The European Union is more than a confederacy. It enjoys some of the powers that normally pertain

only to states” (p. 18). The European Commission is a supranational institution that well exceeds the intergovernmentalism of the Council, and there is a court to adjudicate disputes among the constituent units, as well as between the member states and the union. The European Parliament is the world’s only directly elected multinational legislature, with codecision powers on the rise. Issues of trade and commerce are increasingly the preserve of EU institutions rather than individual nation-states. This conscious coming together of 27 long-established, independent states willing to submit to a higher authority on a wide range of meaningful issues warrants, in this view, the label of federation.

Uncertainty over how to classify the EU and skepticism about its future prompt a further, and especially important, observation, namely, that forms of territorial governance are dynamic and not inevitably successful. Indeed, an additional area of political science research focuses on the failure of federal experiments. Kavalski and Zolkos (2008) hypothesize that federalism will fail if it does not conform to the rules of liberal democracy, if it does not provide an effective institutional setting for the accommodation of identity, and if it is imposed either externally or domestically (by a particular group). Support for the proposition that federalism and authoritarian tendencies are incompatible can be found in the former Yugoslavia, Czechoslovakia, and the Soviet Union. In Yugoslavia from 1943 until 1991 the country functioned on the basis of federal arrangements that linked six republics—Slovenia, Croatia, Bosnia and Herzegovina, Serbia, Montenegro, and Macedonia—whose citizens had a painful and protracted history of interethnic conflict. If Yugoslav federalism aspired to balancing self-rule with shared rule to pacify that conflict and forge a Marxist-socialist consciousness, then in practice it nevertheless rejected political pluralism, monopolized power in the hands of Communists, and strictly limited the rights of individual citizens. The demise and breakup of federal Yugoslavia unleashed civil war and ethnic cleansing. Similarly, in Czechoslovakia from 1918 until its breakup in 1992, liberal democratic values and rights were constrained by Communist rule. When openness came—and came swiftly in the late 1980s—the institutions of Czechoslovakia’s federal state proved incapable of avoiding guilt by association with an illiberal regime and therefore failed to survive the transition and consolidation into post-Communism. Federalism likewise could not hold together the disparate peoples of the Soviet Union. Instead, federalist structures from 1917 through 1991 were instruments of an ideology that sought to control and suppress individual rights and identities rather than to protect them in peaceful coexistence.

Federal systems that fail to accommodate group identities through stable institutional arrangements are also likely to arrive at a premature demise. Examples abound, including those in Ethiopia–Eritrea, Indonesia, and Cameroon. In postcolonial Ethiopia from 1952 to 1962,

United Nations–mandated efforts to incorporate Eritrea employed federalism as a means of strengthening state elites rather than as a mechanism for ethnic accommodation. Elaborate protections of Eritrean identity and autonomy formalized in a constitution were, however, systematically ignored by the Ethiopian government. Policies of assimilation repressed rather than respected Eritrean history and culture, as did the suppression of symbols such as Eritrea’s flag and restrictions on languages. The federation dissolved in 1962, was replaced by annexation, and witnessed 30 years of civil war that culminated in Eritrea’s independence in 1993. A federal system established in postcolonial Cameroon in 1961 failed to achieve institutional balance for that artificial country’s anglophone and francophone communities; instead, the asymmetrical nature of that federal design clearly favored the francophone elites, who orchestrated the federation’s demise in 1972. Cameroon’s defunct federalism made way for a new unitary state structure, rather than outright separation. Other examples of fleeting federations that failed to succeed include Indonesia, which lasted a mere 9 months in 1950. There the advocates of federalism chose a system that sought to impose a national identity from above instead of skillfully protecting and balancing competing identity claims from below.

Empirical evidence also suggests that the success of federalism is less likely when federalism is imposed rather than chosen. British imposition of a federal framework for the West Indies from 1958 to 1962 met with failure and collapse. Similarly, efforts from 1947 to 1971 to establish a federation connecting Pakistan and what was then known as East Pakistan were neither negotiated nor homegrown; as a result, the absence of popular or elite support led to schism, breakup, and the creation of a separate state (Bangladesh). The aforementioned examples of failed federations in Ethiopia–Eritrea as well as the Soviet satellite states provide further support for an inverse relationship between imposition of federalism and its successful implementation.

Policy Implications

If history is littered with examples of territorial governance structures that failed, then it offers a cautionary tale to those reforming old systems or seeking to create new ones in the aftermath of civil conflict or interstate wars. For this reason, political science research on unitary, federal, and confederal arrangements is particularly salient to policymakers seeking solutions to ethnic or territorial discord in places such as Iraq, Afghanistan, and Serbia. There they must ask and answer questions about the institutional forms and procedural safeguards that will give the system the greatest chance to achieve policy objectives (e.g., stability, legitimacy, protection of minorities, production of public goods, and economic efficiencies). Bermeo (2002)

argues that “no violent separatist movement has ever succeeded in a federal democracy” (p. 108). If that is the case, then policymakers in divided societies will have to balance the presumed peaceful by-products of federalism against the traps into which previous but now-defunct federations have fallen.

One of the most contentious policy areas is language. In unitary states, the central government will typically have exclusive jurisdiction over official language rights. That kind of exclusive control frequently coincides with conscious policies of assimilation. By contrast, federal states will normally grant shared or overlapping powers to central as well as regional governments. If federalist principles are pursued, then the ideal policy outcome would be some effective balance between unity and diversity such that citizens using majority and minority languages are equally protected. In Canada, for example, francophone minorities outside Quebec and anglophone minorities inside Quebec enjoy equivalent constitutional protections. The Belgian federation grants exclusive powers to community institutions to regulate language (as well as culture and education). In Spain, which embodies certain elements of federalism without formally constituting a federation, language laws protect Catalonia’s unique language and even make prohibitions against excessive use of the nationally dominant Castilian. India’s federal system provides official status for many of its minority languages, such as Punjabi and Marathi. In transitioning and contested states such as Iraq, policy uncertainty over minority languages could potentially thwart consolidation of a new polity. Recognized alongside Arabic under a 2005 Iraqi constitution, the Kurdish language presently enjoys formal protection and guarantees that elude Kurds in neighboring Iran, Turkey, and Syria.

There are also related policy implications in the area of education. In unitary France, for example, education is so highly centralized that it has long been said that one could look at the clock at any given time and know exactly what textbook schoolchildren throughout the country were reading at that moment. In Canada, by contrast, the federal government has no central education ministry; instead, each province has its own political and administrative infrastructure with discretionary powers over education. Likewise, in the U.S. federal system, education is the responsibility of local (usually county) and state authorities. Pressures to democratize more tenuous federations have led, in places such as Brazil, to policies of educational decentralization. Elsewhere, as in federal Argentina, education has been decentralized to the local level largely on the basis of cost efficiencies and reductions in central government fiscal deficits. Decentralizing and federalizing education can be highly contentious and, quite often, inefficient, as exemplified by the Belgian experience of dismantling universities (including their impressive library collections) and dividing them into separate francophone and Flemish institutions.

Future Directions

Political science research on unitary, federal, and confederal forms of territorial governance has been preoccupied, first, by the struggle to draw clear distinctions between the three concepts; second, by the desire to explain why different institutional designs have been adopted and adapted; and third, by efforts to isolate the impact of federalism on such key outcomes as ethnic group accommodation and system stability. Real advances in our understanding of state forms have emerged, but definitive findings are few and far between. This leaves much room for future research.

Future research would be well served if it followed Eaton’s (2008) advice: “Federalism is an institutional phenomenon, but it should be examined in ways that are sensitive to the importance of noninstitutional variables” (p. 696). In this way, formal-legal studies must be complemented by those that focus at an individual or group level on interests and identities. How is the federal spirit cultivated—or not—at the grassroots level? How are claims for regional autonomy manipulated by political elites seeking to convince groups to join them in a push for system change? How do citizens in systems of compounded representation weigh the competing claims on their identity from local, regional, national, and even supranational governments? What system types yield the greatest citizen satisfaction and loyalty? Do unitary systems insulate their citizens from the “voting fatigue” that seems to bedevil some federations (resulting, for example, in comparatively low voter turnout in the United States and Switzerland)? Even at the institutional level of analysis, future work will need to expand beyond *de jure* constitutional forms to include studies of how those forms interact with *de facto* executive-legislative relations, rules of electoral contestation, party system development, and the importance of informal networks of decision making.

One especially important direction for future research should be that which examines the impact of unitary, federal, and confederal models on democratic accountability. This fundamental element in the debate over representation and territorial governance is presently underinvestigated. Scholars have proffered compelling arguments to suggest that granting voice to regional interests in constitutionally protected upper legislative chambers helps hold central authorities to account for their policies, that empowering regional governments with important functional responsibilities helps bring government closer to the people, and that having meaningful regional institutions helps prepare (and even screen out) politicians for future national office. However, federalism may also introduce unnecessary ambiguity to the process of assigning credit and blame in a political system through overlapping institutions and shared jurisdictions. It may also reduce voter turnout through a proliferation of elected institutions demanding judgments from voters on a nonstop electoral

calendar. Further, regional governments in federal systems may actually represent a form of unintended centralization if those units usurp functions traditionally reserved for local and communal institutions. It remains unclear, then, whether federalism is superior to unitary forms in this crucial dimension of accountability.

Conclusion

The study of territorial dispersion of power in a system of governance is as old as modern political science, and it clearly dates further back to classical debates about the ideal polity. Contemporary efforts to understand the causes and consequences of different state designs have had to struggle first with establishing distinguishable concepts that accurately capture global variation. Scholars have traditionally had a relatively easy time defining unitary systems and classifying states within that category; however, devolutionary trends and reforms within unitary states have created the odd hybrid of decentralized but still unitary polities that defy simple classification. Although numerous attempts have been made to identify the necessary and sufficient conditions for a country to be labeled a federation, it is simultaneously apparent that no two federations are alike. Indeed, the institutional forms that federations take are a function of alternative conceptions of federalism; thus, U.S. and Canadian federal forms are substantially different, as are those in India, Brazil, Australia, and Germany, because the ideological underpinnings of federalism in each of those cases are products of varying historical experiences. Scholars likewise debate the empirical application of confederal approaches to territorial governance, given that some of the clearest cases of confederalism in practice are historical ones that no longer exist. The EU as an exemplar of confederation is only partially appropriate, and the future evolution of European integration will help clarify whether that community of states retains primary power in the hands of its members or whether, instead, it transfers more to an empowered central government to warrant classification as a federation.

Beyond concerns over conceptual boundaries, political science has focused on identifying the factors that facilitate successful negotiation of a federal bargain as well as the conditions that increase the likelihood that reforms of territorial governance will succeed and endure. Such work has drawn attention to cultural predispositions (presence or absence of a federal spirit), the nature of political party systems, elite motives, external intervention and imposition, mechanisms for the accommodation of group identities, and the inverse relationship between successful federalism and illiberal governance. Additionally, the choice of federal, unitary, or confederal forms can be shown to have practical policy implications in such important areas as language laws and education systems. Debate over these choices and their consequences permeates

political discourse in established democracies, in countries that have transitioned from authoritarian or colonial rule, and in systems emerging from civil conflict or interstate war. As such, the contemporary relevance of research on the political and administrative distribution of power cannot be overstated.

References and Further Readings

- Anderson, G. (2008). *Federalism: An introduction*. Oxford, UK: Oxford University Press.
- Bache, I., & Flinders, M. (2004). *Multi level governance*. Oxford, UK: Oxford University Press.
- Bermeo, N. (2002). The import of institutions. *Journal of Democracy*, 13(2), 96–110.
- Burgess, M. (2006). *Comparative federalism: Theory and practice*. London: Routledge.
- Burgess, M., & Gagnon, A. (1993). *Comparative federalism and federation: Competing traditions and future directions*. Toronto, ON, Canada: University of Toronto Press.
- Church, C., & Dardanelli, P. (2005). The dynamics of confederalism and federalism: Comparing Switzerland and the EU. *Regional & Federal Studies*, 15(2), 163–185.
- Deschouwer, K. (2005). Kingdom of Belgium. In J. Kincaid & G. Tarr (Eds.), *Constitutional origins, structure, and change in federal countries* (pp. 48–75). Montreal, QC, Canada: McGill Queens University Press.
- Downs, W. (1999). Accountability payoffs in federal systems? Competing logics and evidence from Europe's newest federation. *Publius: The Journal of Federalism*, 29(1), 87–110.
- Duchacek, I. D. (1970). *Comparative federalism: The territorial dimension of politics*. New York: Holt, Rinehart & Winston.
- Eaton, K. (2008). Federalism in Europe and Latin America: Conceptualization, causes, and consequences. *World Politics*, 60(4), 665–698.
- Elazar, D. (1987). *Exploring federalism*. Tuscaloosa: University of Alabama Press.
- Filippov, M., Ordeshook, P. C., & Shvetsova, O. (2004). *Designing federalism: A theory of self-sustainable federal institutions*. Cambridge, UK: Cambridge University Press.
- Forsyth, M. (1981). *Unions of states: The theory and practice of confederation*. Leicester, UK: Leicester University Press.
- Friedrich, C. J. (1968). *Trends of federalism in theory and practice*. New York: Praeger.
- Gerrig, J., Thacker, S., & Moreno, C. (2005). Centripetal democratic governance: A theory and global inquiry. *American Political Science Review*, 99(4), 567–581.
- Gerston, L. (2007). *American federalism: A concise introduction*. Armonk, NY: M. E. Sharpe.
- Gibson, E. L. (2004). *Federalism and democracy in Latin America*. Baltimore: Johns Hopkins University Press.
- Hueglin, T. (2003). Federalism at the crossroads: Old meanings, new significance. *Canadian Journal of Political Science*, 36(2), 275–294.
- Hueglin, T. O., & Fenna, A. (2006). *Comparative federalism: A systematic inquiry*. Toronto, ON, Canada: University of Toronto Press.

- Karmis, D., & Norman, W. (Eds.). (2005). *Theories of federalism: A reader*. New York: Palgrave Macmillan.
- Kavalski, E., & Zolkos, M. (Eds.). (2008). *Defunct federalisms: Critical perspectives on federal failure*. Aldershot, UK: Ashgate.
- Kincaid, J., & Tarr, G. A. (Eds.). (2005). *Constitutional origins, structure, and change in federal countries*. Montreal, QC, Canada: McGill Queen's University Press.
- King, P. (1982). *Federalism and federation*. Baltimore: Johns Hopkins University Press.
- Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty six countries*. New Haven, CT: Yale University Press.
- Lister, F. K. (1996). *The European Union, the United Nations, and the revival of confederal governance*. Santa Barbara, CA: Greenwood Press.
- Majone, G. (2006). Federation, confederation, and mixed government: A EU US comparison. In A. Menon & M. Schain (Eds.), *Comparative federalism: The European Union and the United States in comparative perspective* (pp. 121-148). Oxford, UK: Oxford University Press.
- Menon, A., & Schain, M. (2006). *Comparative federalism: The European Union and the United States in comparative perspective*. Oxford, UK: Oxford University Press.
- Page, E. C., & Goldsmith, M. J. (1989). *Central and local government relations: Comparative analysis of West European unitary states*. London: Sage.
- Putnam, R. (1994). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Rhodes, R. A. W., Binder, S. A., & Rockman, B. A. (2008). *The Oxford handbook of political institutions*. New York: Oxford University Press.
- Riker, W. H. (1964). *Federalism: Origin, operation, significance*. Boston: Little, Brown.
- Smith, B. C. (1985). *Decentralization: The territorial dimension of the state*. London: Unwin Hyman.
- Steytler, N. (2005). Republic of South Africa. In J. Kincaid & G. Tarr (Eds.), *Constitutional origins, structure, and change in federal countries* (pp. 311-346). Montreal, QC, Canada: McGill Queens University Press.
- Tarr, G., Williams, R., & Marko, J. (Eds.). (2004). *Federalism, subnational constitutions, and minority rights*. Westport, CT: Praeger.
- Tiersky, R. (2002). *François Mitterrand: A very French president*. New York: Rowman & Littlefield.
- Verney, D. (2002). Choosing a federal form of governance for Europe. In A. Heinemann Grüder (Ed.), *Federalism doomed? European federalism between integration and separation* (pp. 17-37). New York: Berghahn Books.
- Watts, R. L. (2008). *Comparing federal systems* (3rd ed.). Montreal, QC, Canada: McGill Queen's University Press.
- Wrede, M. (2004). Small states, large unitary states and federations. *Public Choice*, 119(1-2), 219-240.

21

PRESIDENTIALISM VERSUS PARLIAMENTARISM

JOSEPH W. ROBBINS

Shepherd University

Whether a country employs a presidential or parliamentary regime is an incredibly meaningful distinction. Numerous studies have shown that presidential systems result in more open trade policies and greater particularistic spending (i.e., transportation funding, agricultural subsidies, etc.) and are better suited to represent the entire electorate (Cheibub, 2006; Evans, 2004; Keech & Pak, 1995; Shugart & Carey, 1992). In addition, and perhaps more important, researchers have argued that presidential regimes were prone to conflict and, in some cases, democratic collapse (Linz, 1990a, 1994). More recent work has gone even further to help us understand the nuances of regime type. Margit Tavits's (2009) book considers whether directly elected presidents in parliamentary governments produce greater political divisiveness, pervasive apathy among the electorate, and more frequent intergovernmental conflict. Likewise, much has been written on the benefits of parliamentary governments because they are perceived to be fundamentally different from presidential systems. Specifically, parliamentary governments are thought to engender greater public goods spending (e.g., education, health care, and pensions) and be more efficient and more durable than alternative regime types are. Therefore, it is necessary to highlight the differences between these regimes by focusing on how the chief executive and legislature are elected, how these two branches interact with

one another, and how government formation occurs. Once we can clearly distinguish among alternative government types, we can better identify the relevance for understanding topics such as policy making, representation, and democratic survival.

After this chapter describes the key attributes of presidential and parliamentary systems, it discusses the positive attributes of each system. Furthermore, it addresses the recent proliferation of mixed or hybrid regimes—which contain elements of both presidential and parliamentary systems. Political scientists have long debated the potential dangers or “perils of presidentialism” (Linz, 1990a), and a recapitulation of this discussion represents the final section of this chapter. Finally, the chapter briefly delineates the policy impacts of each regime type.

Theory: Identifying the Main Components of Presidential Regimes

When one thinks of presidential systems, the United States and Latin America typically come to mind because they house a vast majority of all presidential regimes currently in operation. But what exactly is included in a presidential state? Fundamentally, presidential regimes are those in which there is a singularly elected politician who

represents an entire country and whose tenure is not predicated on legislative support. There are also states such as Germany, Hungary, and India that have presidents but whose leaders are comparatively minor players in policy making. Accordingly, the mere existence of a president does not necessarily indicate a presidential state; indeed, the German, Hungarian, and Indian regimes are seldom considered true presidential regimes. Clearly, then, there can be substantial variation among presidential regimes since the aforementioned powers are not possessed by all presidents. In fact, presidents may possess any combination of the following tools: appointment powers, cabinet meeting control, veto power, line-item veto power, emergency powers, foreign policy control, authority over government formation, and the power to dissolve the legislature. Aside from this litany of presidential powers, there is another way to discuss or categorize these systems. That said, perhaps it is more useful to focus on four key characteristics of presidential regimes that can be used to differentiate among executive types.

In Matthew Shugart and John Carey's (1992) impressive book, *Presidents and Assemblies*, the authors discuss the relationship between regime type and electoral system and the subsequent implications for how the two concepts interact. Aside from this discussion, Shugart and Carey focus on a number of key characteristics that can help students classify governments as presidential, parliamentary, or mixed. In particular, they argue that there is an inherent trade-off between presidential and parliamentary governments: Presidential governments tend to be more efficient than parliamentary regimes, but the former may not be as representative as the latter. Efficiency, in this case, refers to voters' ability to identify the final government composition in advance. That is, a presidential system is efficient in that voters can identify who will serve as president, or lead the government, prior to the casting of the final vote. In the 2008 presidential election, voters knew that either Senator John McCain or Barack Obama would become the 44th president of the United States. In contrast, parliamentary systems are not as efficient since the final government composition is contingent on the final electoral results and the subsequent political wrangling over the formation of a governing coalition. On the other hand, representation or representativeness refers to how much of the population is represented by the ruling government. Although presidents can theoretically represent the entire populace, this is certainly not a given. The oft-cited example of Salvador Allende's 1970 electoral victory in Chile, with just over 30% of the popular vote, serves as a reminder that presidents seldom represent the entire populace. In parliamentary systems, moreover, representation of the entire electorate is more easily attained through coalitional governments than under a single leader in a presidential system. Thus, parliamentary regimes offer greater representation but are typically less efficient than their presidential counterparts.

Beyond this demarcation, Shugart and Carey (1992) reason that there are other certain essential attributes found in all presidential systems. First, they claim that a separation of powers among branches of government is necessary. For students of American politics, this concept is well-known. A separation of powers refers to a clear division of responsibility, in which most often the executive branch administers the law, the legislature writes the laws, and the judiciary interprets or reviews the constitutionality of the laws. In parliamentary governments, however, such a division of authority is often lacking, as the legislative and executive branches are essentially fused together.

Second, Shugart and Carey (1992) note that presidents are directly elected through some type of nationwide vote. Their ascension into office is not contingent on parliamentary support, but rather they are elected by an entire country. Of course, this too can take place through various mechanisms. Some states use a simple-majority voting system, in which the winner secures most (or a *plurality*) of the popular vote. Meanwhile, countries such as France use a slightly modified format, where the winner must possess more than 50% of the nationwide vote. In those elections in which there is no clear-cut majority winner, the top two vote getters will compete against each other in a second, runoff election. Another approach uses an indirect mechanism, best illustrated by some type of *electoral college*, to elect the president. Although it is not the only country to use this tool, the United States is perhaps the most well-known country to use this complex mechanism. In the case of the United States, both the Republican and the Democratic Party select a slate of electors for each state equal to the number of House of Representatives members and senators from that state. The presidential candidate who wins a plurality of a state's popular vote receives the Electoral College votes of that same state. Each candidate's electoral votes are then summed, and the candidate with a majority of the Electoral College votes (i.e., 270) is declared the victor.

Third, presidents serve fixed terms that are not dependent on the support or confidence of the legislature. Presidents, in most cases, serve 4- or 5-year terms, and at the end of a term, the president must either step down from the office or run for reelection. Put differently, presidents are not permitted to remain in office indefinitely, whereas (theoretically) prime ministers in some parliamentary governments are. It is worth noting, however, that certain provisions are in place in the event an official commits an egregious offense. For example, following the Bill Clinton–Monica Lewinsky scandal during the 1990s, the U.S. Congress impeached the president, although it did not remove him from office. Similarly, during the First Russian Republic, the Russian legislature sought to impeach Boris Yeltsin for his disastrous economic policies, the First Chechen War, and a host of other calamities. Thus, in most cases, there are mechanisms in place to remove from office a much maligned president, but in ordinary circumstances the president's term will conclude

after 4 or 5 years in most cases. And, unlike in parliamentary systems, the president's term is not susceptible to legislative approval. In parliamentary governments, a prime minister may be asked, or forced, to step down if he or she loses the confidence or support of his or her legislative counterparts. In contrast, a president can be eternally at odds with legislators and not jeopardize his or her political survival. Therefore, a presidential regime is also characterized by fixed terms, which are not contingent on legislative approval. This design is popular for numerous reasons, but it is not without its critics. These topics, though, will be elucidated later in this chapter.

That presidents retain the power to form their own cabinet is the fourth component of presidential structures. Having the authority to appoint cabinet members is a crucial tool at the president's disposal, and this can help the president pursue, and, it is hoped, fulfill, his or her campaign promises. Without this power, a president may have to resort to political negotiating with rivals to fill these positions. In such a scenario, a president's appointments may be delayed as a result of political wrangling between parties or candidates. Such occurrences may hinder a president's ability to carry out his or her mandate, and such a failure could have economic, social, or political repercussions. After all, if a president struggles to appoint defense, education, or economic ministers, then these agencies' ability to carry out changes in defense, education, or economic policies could be undermined. This authority is significantly different under parliamentary regimes, in which cabinet appointments are typically more of a cooperative endeavor.

The final primary attribute in presidential systems involves the executive's lawmaking authority. In some states, presidents have virtually no lawmaking authority. The U.S. president, for example, has some tools available to circumvent Congress's policy-making authority. He can issue executive decrees and executive treaties that can, but do not always, carry the same weight as congressional legislation. In other countries, however, presidential lawmaking authority rivals that of the legislature. As a result, some legislative authority and/or autonomy may be usurped by an overly ambitious executive. One such case is the Russian and Ukrainian presidencies, in which presidents have wielded considerable power (Protsyk, 2004). In contrast, the French executive is considerably different because of the accepted practice of *cohabitation*. During cohabitation, in which the legislature and executive are controlled by different political parties, the president often allows the prime minister to lead the government. This scenario can also include a president's *veto power*, whereby the executive can prevent a bill from becoming a law.

Appeals and Implications of Presidentialism

Before turning to the potential dangers of presidential governance, it is necessary to identify the positive attributes of these systems. To this extent, those who study regime

types have emphasized three primary aspects of presidential states that help us understand their appeal to politicians throughout the world. The three appeals include efficiency in political results and governance, the representativeness of the entire populace, and the checks and balances typically found in these governments.

As discussed earlier, efficiency is an attractive aspect of most presidential systems because of the information conveyed to voters. In particular, efficiency means that voters know, prior to casting a ballot, what the new government will look like. That is not to say that the outcomes are pre-ordained but rather that the most likely scenarios are widely known. This efficiency provides voters with a wealth of information as they decide which candidate to support. They also know that, if elected, their candidate should pursue with few obstacles or restrictions the policies supported by his or her electorate. Parliamentary systems, however, differ considerably. In some instances, the described scenario can, and does, unfold in parliamentary regimes. More common, however, is the formation of a coalition government. In such a situation, two or more parties join together in a coalition. Consequently, each party's demands or expectations are tempered as a result of the coalition formation exercise. Therefore, if a voter supported a conservative party, then the voter's preferred policies may not come to fruition because parties must compromise their positions in order to join the governing coalition. Furthermore, a singularly elected official (e.g., a president) may make it easier for voters to hold government officials accountable. At times, the president may be assigned too much blame if economic problems arise, but it is generally easier to punish one official than an entire legislature for lackluster economic policies. Therefore, efficiency is more apparent in presidential systems in comparison with parliamentary ones.

The second positive contribution of presidential regimes is the representativeness of one elected official. In parliamentary systems, with multiple political parties, it is likely that each party will represent only its supporters. For conservative voters, this means that their needs will be ignored if the majority winner is a liberal party that seeks to represent (or reward) its supporters. Such a possibility becomes even more complex if we consider a country with several political parties. Supporters of presidential governments, however, contend that this scenario is less likely to occur in a presidential state. They argue that presidents represent the entire country and, therefore, are beholden to the entire electorate rather than just a subset of the population. In 2000, when George W. Bush was declared president of the United States, he stated that he would represent the entire country rather than just his supporters, despite his narrow margin of victory. And 8 years later, Barack Obama made a similar pledge as president. Accordingly, a single elected official can, at least theoretically, be better positioned to represent an entire country than may be likely under parliamentary governments. Whether this happens is certainly debatable.

Presidential systems may also be preferred to parliamentary regimes because a presidential system balances representation with another branch of government. That is, most presidential regimes distribute powers among the branches of government. This arrangement provides multiple points for citizens to influence their government. If lobbying the president proves fruitless, then a concerned citizen can turn his or her attention to the legislature. Thus, states with popularly elected presidents can balance or distribute power over multiple branches of government. This institutional fragmentation can also prevent one branch from becoming too powerful and thus running roughshod over minority interests. Therefore, people from all ideological backgrounds should have some say in policy making as long as there are multiple political actors or institutions with some semblance of power. Because the electorate is represented by at least two different institutions (i.e., the legislature and the executive in this case), presidential systems can foment democratic stability by offering voters more opportunities to influence the policy-making process.

Theory: Identifying the Main Components of Parliamentary Regimes

Although presidential systems are common in the Western hemisphere, they are not the only option available to governments. In fact, presidential systems are in the minority when it comes to regime types found throughout the world. Parliamentary systems actually outnumber presidential states when we take stock of the entire global community. To many students of U.S. politics, parliamentary systems remain a foreign concept, and the inner workings of these regimes are equally befuddling. Accordingly, it is necessary to discuss the defining characteristics and operation of parliamentary governments before presenting the appeals of parliamentary systems and discussing an emerging type of government, generally referred to as *hybrid regimes*.

Often referred to as the *Westminster model*, after the United Kingdom's government, parliamentary systems differ from presidential states in several ways. One of the key defining characteristics of parliamentary systems is the fusion of the executive and legislative branches. Whereas the heads of state and government are often embodied in the same person in a presidential regime, parliamentary systems often separate the two roles. In addition, presidential systems generally have a separation of powers among the various branches of government, but such a clear allocation of responsibility is not found in most parliamentary regimes. Instead, parliamentary governments often combine the responsibilities of both the legislative and the executive branches.

After an initial election, in which voters decide how many seats are allocated to the various political parties, the elected representatives in a parliamentary system are then given the task of establishing or forming the government. Not only do these individuals have to organize the

legislative branch, but they are also charged with structuring the executive branch. Legislative officials determine who will serve as the head of the government (or who will serve as the prime minister or premier), which politicians will fill the various cabinet positions, and who will head the various legislative committees. If there is a clear majority winner in the initial elections, then government formation is fairly straightforward, and typically no coalition is formed. In other instances, however, where there is no majority party, the party with the most seats (often referred to as the *formateur* party) is responsible for constructing the governing coalition. This party will seek a coalition partner—or sometimes multiple partners—and, after cobbling together enough support to give the coalition a majority of seats, the parties involved will then jointly determine the prime minister, cabinet positions, and other leadership posts. Although this is the norm, there are cases when a coalition government is not formed, even in the absence of a majority winner. Currently, the Canadian government has a minority government led by the Conservative Party. In the Canadian case, there is no majority party, and coalition formation efforts have proved futile. Consequently, the Conservative Party, by default, is the *de facto* governing party even though it lacks a clear majority of seats.

Cabinet dominance is another important feature of parliamentary democracies (Lijphart, 1999). On forming the ruling coalition, those in power make a number of key appointments, although none more powerful than that of the prime minister. The prime minister is enlisted to serve at the behest of the governing majority and will do so until he or she loses the support of the legislature. When the prime minister loses support of the legislature, as evidenced through a *vote of no confidence*, then new elections must be held. An alternative practice is the *vote of confidence*, which is an act initiated by the government. Here, if the ruling government is incapable of securing a majority of the votes in the legislature, then the government must step down (Clark, Golder, & Golder, 2009). Some states use a slightly different version of the confidence vote, referred to as a *constructive vote of confidence*. In Germany and Hungary, two countries where such a tool is in use, the legislature must agree on a replacement government prior to dissolving the extant body.

Predictably, cooperation among politicians is very important in parliamentary systems, for without it, all elected officials must run for reelection and thus risk losing their positions. Lacking party discipline or cooperation, the government is likely to collapse. This is exactly what has happened recently in the Czech Republic. Like so many other countries in the past few years, the Czech Republic was plagued by the recent global economic downturn, and when economic troubles combined with internal political strife, it was hardly surprising that the governing coalition lost its grasp on power.

Although coalitional survival is predicated on party discipline and cooperation, the most prominent or powerful

actor in most parliamentary governments is the prime minister. Prime ministers are elected by their governing coalition; but not all prime ministers are equally powerful. Indeed, Giovanni Sartori (1994) explains that there are at least three scenarios common to most parliamentary governments: A prime minister may be *first above unequals*, *first among unequals*, or *first among equals*. The power of the prime minister is greatest in the first case (i.e., first above unequals), examples of which can be found in places such as Germany, Greece, and the United Kingdom (Lijphart, 1999). In contrast, among the weakest prime ministers (i.e., first among equals) are the heads of government in Italy, the Netherlands, and Norway (Lijphart, 1999). How does one discern between powerful and weak prime ministers? Typically, researchers have relied on the prime minister's authority vis-à-vis fellow members of the executive branch (e.g., the Exchequer in the United Kingdom compared with the prime minister), the prime minister's ability to navigate through the policy-making process, and his or her ability to remove and appoint members of the executive branch. These are but a few examples of how we can measure prime ministerial strength; many more examples can be found in other sources (see King, 1994; Lijphart, 1999).

Beyond these features, what other defining characteristics do we see in parliamentary governments? Other than the aforementioned lack of separation between the legislature and executive, a few additional attributes are worth stressing. In particular, parliamentary governments are often more conducive to cooperation. Because their political survival depends on cooperation, members of the legislature are more willing to work with the executive branch than are legislatures in presidential regimes. In addition, although established and fully functioning judiciaries are common in many democracies, the separation of powers or checks and balances between the legislative and executive branches found in many presidential governments is often missing from parliamentary governments.

In most parliamentary governments, we also see the use of a *proportional representation* (PR) electoral system to translate votes into legislative seats. There are some exceptions to this. Great Britain and India each employ a *single member district* (SMD) plurality electoral system to fill legislative seats. The use of the SMD electoral system in Great Britain has largely favored the two major parties—the Conservatives and the Labour Party—at the expense of the Liberal Democrats. Although there are other examples of parliamentary systems that use SMD electoral systems, more often than not parliamentary systems employ a PR system. Under these electoral systems, we typically see the emergence of multiple-party systems because more seats per electoral district are up for grabs. Maurice Duverger (1954), a noted French political scientist, observed many years ago that PR systems typically produce multiparty systems while SMD systems typically have a constraining effect on the number of political parties. More specifically, he explained that most SMD systems should produce

two-party systems like those seen in the United States and, to a similar degree, the United Kingdom.

Multiple-party systems are often found in parliamentary governments that use PR electoral systems. This can be especially attractive to highly fragmented societies. In fact, given the deep sectarian divides in Iraq, it should not be surprising that the architects of this young democracy opted for a parliamentary style of governance with a PR electoral system. This structure enables Iraqi leaders to represent a multitude of diverse interests—including political, ethnic, religious, and geographic. The choice of government type and electoral system is a significant decision; indeed, this setup should allow multiple parties to participate in the policy-making process, which should, at least in theory, temper Iraqi tensions. That parliamentary governments with PR electoral systems are better able to represent multiple interests is one of the more attractive features of this regime type.

Another distinguishing feature of parliamentary governments is the absence of fixed terms. The confidence votes used to sustain a coalition government suggest that a government can fail at any time. Although most parliamentary states have provisions that require elections to be held every 4 or 5 years, virtually all parliamentary terms can be cut short. Elections that may be held at any time are also referred to as *endogenously timed elections*. Recent political science research has found that politicians may use this feature to extend their political careers: When a government is viewed in a positive manner, elections may be called earlier to capitalize on this success, but when public approval ratings are low or decreasing, then the governing coalition will often postpone elections. Others, however, have found that since political leaders possess more accurate information than voters, early elections will be held if elected officials anticipate political strife or economic turmoil (Smith, 2003). As a corollary, other works have found that voters are cognizant of this practice and, in some cases, may punish coalition members for holding elections too early. Nevertheless, this is an attractive feature of the Westminster model as it gives the electorate more tools to keep tabs on elected officials.

Appeals and Implications of Parliamentary Governments

Parliamentary systems offer several remedies to the majoritarian tendencies common to most presidential governments. For starters, many suggest that accountability is in fact greater in parliamentary—not presidential—systems. While ineffective presidents could scapegoat their rivals in the legislature, this argument is unconvincing for leaders in parliamentary systems. Because there is no other actor who can thwart policy making, coalition members are more accountable for the results of their policies. As one can imagine, this is a double-edged sword. On one hand, if the government's policies are utterly disastrous, then everyone knows whom to blame. Thus, it is easy to

remove from power those responsible for the ineffective or harmful policies. On the other hand, if a government's policies prove to be a rousing success, then this can certainly work to the coalition members' advantage. Because it was the coalition's steady hand alone that implemented the policies responsible for economic growth, political stability, and so forth, its members are poised to be the lone beneficiaries of the policies in question. As such, advocates of the parliamentary model would suggest that this style of governance is better able to reward or punish the politicians at the helm of the state apparatus.

Similarly, the presence of endogenously timed elections is also seen as a boon for accountability and, by extension, democratic stability. This mechanism rewards popular or highly regarded politicians for their policies by granting them an extended tenure. One of the arguments against presidentialism is that if a president has done an excellent job, he or she still must run for reelection at the end of a term. But in parliamentary systems, if those in power are highly regarded, then there is no need to hold elections (again, there are exceptions to this). Equally telling, though, is the constant threat of dissolving the government. If a country's political or economic climate has suffered at the hands of a coalition government, then that government can be removed from office at any time. This enables voters to oust unpopular politicians at any time, which is an improvement over presidential structures. Barring any significant scandal or an attempted coup d'état, most presidents cannot be removed from office until the expiration of their term. Consequently, a nefarious or incompetent president cannot be removed prematurely (unless impeachment proceedings are initiated), thus enabling the president to wreak further havoc on a country's livelihood.

Finally, parliamentary systems are thought to be more conducive to political party formation, especially multi-party systems. While this likelihood is partly attributed to the frequent use of PR electoral systems in parliamentary governments, party systems also thrive under parliamentary governments because of the incentives to create and sustain parties in this type of regime. As Duverger noted long ago, PR contests foment multiparty systems. This notwithstanding, it is also worth noting that presidential systems tend to constrain party-building efforts. After all, if a president is endowed with adequate constitutional power with respect to the legislature, then he or she need not be overly concerned with working with the legislature. Furthermore, when presidents can circumvent the legislature, there is little incentive for legislators to build or maintain political parties. Thus, political parties may be less instrumental in lawmaking under presidential regimes. Conversely, in parliamentary governments, political parties are effective vehicles for overcoming collective action problems that often surface in legislative bodies (Aldrich, 1995). Creating and maintaining parties, moreover, can help legislators overcome collective action problems through organizing members, delivering necessary

information, and rewarding loyal rank-and-file members. By providing these benefits, individual legislators will succumb to the party in hopes of furthering their careers. And, in exchange for this loyalty, parties are better able to legislate given this increased party discipline. Without political parties, parliamentary systems could become immobilized due to the rampant chaos stemming from the lack of organization and structure. Likewise, we can point to the Weimar Republic in Germany and the French Fourth Republic as two examples of what happens when party organizations are incipient or unable to coerce discipline from their rank-and-file members. In both cases, the weak parties populating the political landscape undermined policy-making efforts.

Perils of Presidentialism: Do Presidential Systems Produce Democratic Instability?

The decision of whether to implement a parliamentary or a presidential system has engendered a healthy debate among scholars. Much of this discussion was fueled in part by the experience of Latin America, where many countries implemented presidential systems that would collapse years later. Because of this turbulent regional experience, a number of scholars were quick to write off presidential regimes because, they argued, these systems would undermine democracy and lead to political unrest. The reasons for this skepticism rest on four primary criticisms. Ostensibly, the argument contends that presidentialism hinders democratic survival due to the prevalence of minority governments, frequent legislative impasses, the lack of durable coalitions, and the propensity of the executives to exploit constitutional authority for their own benefit.

At their core, presidential systems are essentially majoritarian, winner-takes-all contests. That is, the second-place finisher receives nothing in return for his or her strong showing. When George W. Bush narrowly edged out Al Gore in the Electoral College in the 2000 U.S. presidential election, there was no consolation prize for then Vice President Gore even though he had in fact won the popular vote. Given this, the winner of a presidential contest has considerable authority and can do as he or she pleases. Furthermore, if the losing candidate is or was serving in the legislature, then the victorious candidate has little incentive to work with his or her former opponent. This situation can result in a political stalemate that could prompt drastic measures that could have a devastating effect on the overall political system.

Furthermore, because there is only one president, presidential systems also tend to produce minority governments. One excellent example of such a scenario is Chile's Allende, who became president after winning only slightly more than 30% of the popular vote. A similar case is the 2002 French presidential elections. Due to pervasive frustration with their government, much of the French electorate voted against

familiar candidates, and nearly 20% of voters supported the right wing candidate Jean-Marie Le Pen—a caustic individual whose campaign has been described as xenophobic if not anti-Semitic in many instances. In France's two-round presidential contest, Le Pen was one of two candidates to make it to the second round, in which he was soundly defeated by Jacques Chirac. Nevertheless, these examples, particularly Allende's story, serve as subtle reminders that presidentialism may result in minority governments. Furthermore, in the case of Le Pen, had he been elected, it is plausible that his policies would have contrasted sharply with the views of much of the French electorate, thus causing a potentially dangerous scenario.

Another reason presidents may hinder democratic survival stems from what is commonly referred to as dual democratic legitimacies. Because presidents and legislators are elected in separate contests—unlike in parliamentary systems—it is possible, if not probable, that both the legislative and executive branches of government may claim separate political mandates. If there are rules or provisions in place to guide the government through these potentially rough waters, then there is little cause for concern. However, in most cases, there is little clarity in doling out responsibilities between these branches of government, in which case, any showdown between the legislature and the president could result in a political stalemate and, possibly, democratic collapse. The latter is certainly not a given, but at the very least, legislative gridlock is likely to ensue. If a political shutdown does not result, the separate mandates of each branch could result in an impasse that could at least stall policy making, which could deprive government agencies, citizens, and politicians of potentially valuable policies or funding that is needed to keep society running smoothly.

Additionally, if the president and legislative majority hold different or contrasting ideological backgrounds or partisan affiliations, then any quasi coalition will likely be short-lived. Admittedly, this is certainly not a given, but it is a likely outcome when the president and legislature hold different views on the direction of a given country. And since there are few incentives for these actors to work together, coalitional stability is likely to be tenuous at best. This instability could slow the political process as these institutions battle for influence, thus resulting in delayed or nonexistent policy making. If policy making does not grind to a halt, then an alternative scenario exists whereby presidents would simply eschew the traditional lawmaking process and opt for extralegal options. This too could precipitate a political showdown that could trigger a democratic crisis or, worse yet, collapse.

Faced with this scenario, many presidents were quick to exploit the constitution for their own personal gains. This potential consequence of presidentialism is based on numerous experiences in Latin America, where presidents attempted to cling to power even when faced with mounting opposition from the legislature and/or the polity. This consequence can be viewed as the culmination of the other

components identified thus far. In the event that a president seeks to rewrite, distort, or even ignore the constitution, the fate of democracy has been sealed, and it will be on the brink of collapse.

Although much has been written about the potential perils of presidentialism, this debate remains ongoing. Other scholars have questioned this contention and have countered that presidential systems are not inherently dangerous. For example, Donald Horowitz (1990) takes umbrage with a number of Juan Linz's conjectures. First, Horowitz points out that Linz's evidence is predominantly based on the Latin American experience and neglects stable presidencies found in other regions. Second, Horowitz maintains that the perils of presidentialism are based primarily on exaggerated interpretations of these systems that do not represent all presidential states. It is important to note that Horowitz also points out several instances in which parliamentary systems have prompted instability. Among others, Nigeria is mentioned as one case in which parliamentarism has resulted in an unstable polity. Another point of contention stems from the confusion surrounding presidential systems and electoral systems. Those subscribing to the perils-of-presidentialism school of thought have essentially confused presidential systems with electoral systems, thus producing some misguided conclusions. For example, the single member plurality systems used to elect many presidents often result in disproportional outcomes that potentially exacerbate underlying social, political, or geographic tensions. Finally, those questioning the dangers of presidentialism also point out that the potentially beneficial aspects of presidential systems have been largely overlooked, which has provided a rather biased view of these regimes.

Although useful, this theoretical discussion does not provide much empirical evidence of the dangers of presidentialism. Fortunately, a number of political scientists have used a variety of statistical techniques to evaluate the perils-of-presidentialism argument. Alfred Stepan and Cindy Skach (1993) have reviewed several constitutional frameworks, along with a host of other data sources, and have discovered that presidentialism is indeed correlated with weaker democracies. Not surprisingly, others have reached much different conclusions. In their study of more than 50 countries, Timothy Power and Mark Gasiorowski's (1997) findings refute the dangers of presidentialism. In the most comprehensive work to date, Jose Antonio Cheibub (2007) has echoed Power and Gasiorowski's conclusions. Cheibub (2007) asserts that it is not presidential systems that are dangerous; rather, he demonstrates that it is the underlying social conditions or background characteristics that jeopardize democratic stability. Specifically, he suggests that whether or not a military regime had previously existed is far more important than whether a presidential system is in place. Because many Latin American countries have had experiences with military dictatorships, Cheibub's findings put Linz's theoretical work in a new perspective.

Future Directions

Clearly, much has been written on the topic of regime choice. Convincing arguments have been made on both sides of the debates—those warning of the perils of presidentialism and those questioning this assertion. Cheibub's work is a compelling study of the presidentialism argument; here, the author disaggregates the various components of the presidentialism argument and finds that statistical evidence dispels many of the concerns espoused by the camp opposed to presidentialism. This work, however, does not address every concern. In fact, Cheibub goes on to say that testing the relationship between political party discipline and presidents is impaired by our inability to measure party discipline cross-nationally. Aside from this empirical problem, there are other concerns with how we measure these regimes. Although scholars such as Alan Siaroff (2003) include several considerations to measure presidential strength, our ability to clearly distinguish presidential states from parliamentary ones is still a work in progress. This issue is further complicated when we take into account the large number of hybrid systems that are also particularly elusive in categorizing regime type.

Conclusion

There are several different ways a country can set up its own government. It can take the form of a presidential, parliamentary, or hybrid regime. To be sure, there are attractive aspects of each regime, but it should not be surprising that many newer democracies have opted for a hybrid regime that seeks the best of both worlds. For example, several former Communist states, such as Croatia, Lithuania, Poland, and Ukraine, have opted for semipresidential regimes, in part to mitigate the shortcomings of either presidential or parliamentary systems (Clark, Golder, & Golder, 2009). Having said that, simply because a country opts for a presidential regime does not necessarily precipitate democratic collapse. For some time, scholars had reasoned that presidential systems were anathema to democracy, but contemporary analysis has largely disputed this finding. Thus, presidential systems may not be as perilous as Linz and others have forecast. This is not to suggest that regime type is insignificant. Rather, regime type remains an important distinction, and it is imperative that we understand the costs and benefits of different regime types in order to minimize the potential negative consequences for a given country.

References and Further Readings

- Aldrich, J. (1995). *Why parties? The origin and transformation of party politics in America*. Chicago: University of Chicago Press.
- Baldez, L., & Carey, J. (2001). Budget procedure and fiscal restraint in posttransition Chile. In S. Haggard & M. D. McCubbins (Eds.), *Presidents, parliaments, and policy* (pp. 105–148). Cambridge, UK: Cambridge University Press.
- Cheibub, J. A. (2002). Minority governments, deadlock situations, and the survival of presidential democracies. *Comparative Political Studies*, 35, 284–312.
- Cheibub, J. A. (2006). Presidentialism, electoral identifiability, and budget balances in democratic systems. *American Political Science Review*, 100, 353–368.
- Cheibub, J. A. (2007). *Presidentialism, parliamentarism, and democracy*. New York: Cambridge University Press.
- Clark, T. D., & Wittrock, J. N. (2005). Presidentialism and the effect of electoral law in postcommunist systems: Regime type matters. *Comparative Political Studies*, 38, 171–188.
- Clark, W. R., Golder, M., & Golder, S. N. (2009). *Principles of comparative politics*. Washington, DC: CQ Press.
- Cox, G. W. (1987). *The efficient secret: The cabinet and the development of political parties in Victorian England*. Cambridge, UK: Cambridge University Press.
- Duverger, M. (1954). *Political parties*. New York: Wiley.
- Evans, D. (2004). *Greasing the wheels: Using pork barrel projects to build majority coalitions in Congress*. New York: Cambridge University Press.
- Frye, T. (1997). A politics of institutional choice: Post Communist presidencies. *Comparative Political Studies*, 30, 523–552.
- Hellman, J. S. (1996). Constitutions and economic reform in the postcommunist transitions. *East European Constitutional Review*, 5, 46–56.
- Horowitz, D. (1990). Comparing democratic systems. *Journal of Democracy*, 1, 783–794.
- Ishiyama, J. T., & Kennedy, R. (2001). Superpresidentialism and political party development in Russia, Ukraine, Armenia and Kyrgyzstan. *Europe Asia Studies*, 53, 1177–1191.
- Ishiyama, J. T., & Velten, M. (1998). Presidential power and democratic development in post Communist politics. *Communist and Post Communist Studies*, 31, 217–233.
- Keech, W. R., & Pak, K. (1995). Partisanship, institutions and change in American trade politics. *Journal of Politics*, 57, 1130–1142.
- King, A. (1994). "Chief executives" in western Europe. In I. Budge & D. McKay (Eds.), *Developing democracy: Comparative research in honor of J. F. P. Blondel* (pp. 150–163). London: Sage.
- Laver, M., & Shepsle, K. A. (1996). *Making and breaking governments: Government formation in parliamentary democracies*. New York: Cambridge University Press.
- Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty six countries*. New Haven, CT: Yale University Press.
- Linz, J. J. (1990a). The perils of presidentialism. *Journal of Democracy*, 1, 51–69.
- Linz, J. J. (1990b). Virtues of parliamentarism. *Journal of Democracy*, 1, 83–91.
- Linz, J. J. (1994). Presidential or parliamentary democracy: Does it make a difference? In J. J. Linz & A. Valenzuela (Eds.), *The failure of presidential democracy* (pp. 3–87). Baltimore: Johns Hopkins University Press.
- Lohmann, S., & O'Halloran, S. (1994). Divided government and U.S. trade policy: Theory and evidence. *International Organization*, 48, 595–632.

- Mainwaring, S. (1993). Presidentialism, multipartism, and democracy: The difficult combination. *Comparative Political Studies*, 26, 198-228.
- Marsteintredet, L., & Berntzen, E. (2008). Reducing the perils of presidentialism in Latin America through presidential interruptions. *Comparative Politics*, 41, 83-102.
- McCarty, N. (2000). Presidential pork: Executive veto power and distributive politics. *American Political Science Review*, 94, 117-129.
- McGregor, J. P. (1994). The presidency in east Central Europe. *RFE/RL Research Report*, 3, 23-31.
- McGregor, J. P. (1996). Constitutional factors in politics in post Communist central and eastern Europe. *Communist and Post Communist Studies*, 29, 147-166.
- Metcalf, L. K. (2000). Measuring presidential power. *Comparative Political Studies*, 33, 660-685.
- Persson, T., Roland, G., & Tabellini, G. (2000). Comparative politics and public finance. *Journal of Political Economy*, 106, 1121-1160.
- Persson, T., & Tabellini, G. (1999). The size and scope of government: Comparative politics and rational politicians. *European Economic Review*, 43, 699-735.
- Power, T. J., & Gasiorowski, M. J. (1997). Institutional design and democratic consolidation in the third world. *Comparative Political Studies*, 30, 123-155.
- Protsyk, O. (2004). Ruling with decrees: Presidential decree making in Russia and Ukraine. *Europe Asia Studies*, 56, 637-660.
- Roper, S. D. (2002). Are all semipresidential regimes the same? A comparison of premier presidential regimes. *Comparative Politics*, 34, 253-272.
- Sartori, G. (1994). Neither presidentialism nor parliamentarism. In J. J. Linz & A. Valenzuela (Eds.), *The failure of presidential democracy* (pp. 106-118). Baltimore: Johns Hopkins University Press.
- Shugart, M. S., & Carey, J. M. (1992). *Presidents and assemblies: Constitutional design and electoral dynamics*. Cambridge, UK: Cambridge University Press.
- Siaroff, A. (2003). Comparative presidencies: The inadequacy of the presidential, semipresidential and parliamentary distinction. *European Journal of Political Research*, 42, 287-312.
- Smith, A. (2003). Election timing in majoritarian parliaments. *British Journal of Political Science*, 33, 397-418.
- Stepan, A., & Skach, C. (1993). Constitutional frameworks and democratic consolidation. *World Politics*, 46, 1-22.
- Tavits, M. (2009). *Presidents with prime ministers: Do direct elections matter?* Oxford, UK: Oxford University Press.

22

COMPARATIVE JUDICIAL POLITICS

DAVID L. WEIDEN

Indiana University–Purdue University Indianapolis

The study of courts in comparative perspective has a long history in the discipline of political science. C. Neal Tate (2002a) notes that articles dealing with international judicial systems were published nearly 100 years ago in some of the first American political science journals. However, the study of comparative judicial politics faded from the forefront of the discipline of political science for several generations, until social scientists began to again analyze legal systems comparatively, starting in the 1960s. In the last several decades, there has been a virtual explosion of research dealing with judicial structures and processes in comparative perspective. This chapter provides an overview of some of the most significant recent research on comparative courts and also examines the trends and future directions for comparative judicial scholarship. Obviously, only the general trends in comparative courts scholarship can be discussed here due to space limitations; for a more detailed listing of work in the field, the reader is advised to consult Tate (2002a, 2002b), which provides a comprehensive bibliography of books and articles in the subfield of comparative judicial systems.

This chapter first reviews the primary theories of judicial decision making, then examines constitutional and appellate courts in comparative perspective, including the differing forms of judicial review found throughout the world, then provides an overview of comparative trial systems, and concludes with suggestions for future directions in comparative judicial politics research.

Courts and Legal Systems

The study of courts in comparative perspective requires that one first determine exactly what constitutes a court. Early work in the comparative study of judicial systems attempted to provide this definition. An important study is Theodore Becker's *Comparative Judicial Politics: The Political Functionings of Courts* (1970). In this study, Becker defined a *court* as follows:

- (1) a man or body of men (2) with power to decide a dispute, (3) before whom the parties or advocates or their surrogates present the facts of a dispute and cite existent, expressed, primary normative principles (in statutes, constitutions, rules, previous cases) that (4) are applied by that man or those men, (5) who believe that they should listen to the presentation of facts and apply such cited normative principles impartially, objectively, or with detachment, . . . and (6) that they may so decide, and (7) as an independent body. (p. 13)

A functionalist definition of *court* is given by Martin Shapiro in his important book *Courts: A Comparative and Political Analysis* (1981). Shapiro suggests that courts serve three major functions: conflict resolution, social control, and lawmaking. By *lawmaking*, Shapiro refers to interstitial lawmaking, which consists of filling in the gaps in statutory or customary law, and also to the judicial creation of law and policy. He notes that, although the existence of

judicial policy making is frequently denied, the phenomenon exists in all courts and judicial systems.

These early works are notable because they laid out the foundations on which more advanced analyses in comparative judicial politics could be conducted. The examination of comparative court systems is complicated by the fact that there are three general types of legal systems found in the world: common law, civil law, and religious law. The jurisprudence, judicial role, and organization of courts vary widely in each of these legal systems. Put most simply, *common law* legal systems are found primarily in Anglo-American nations, and the principle of precedent—following the legal principle established in previous court cases—is one of the chief elements of this legal system. *Civil law* systems, found in Continental Europe, Latin America, and much of the rest of the world, generally do not rely on the principle of precedent but instead use extensive codification and written codes to guide judges in their decision making. Finally, the great variety of *religious law* systems, found in the Middle East and elsewhere, rely extensively on the principles found in sacred texts to direct judges in their duties (see, generally, Merryman, 1985).

Clearly, the wide variance in the types of courts and legal systems has presented significant obstacles for researchers seeking to understand and elucidate legal phenomena across national borders. Thus, the most common type of research in comparative judicial research is the single-country or single-court study, which seeks to explain and predict judicial behavior in one court only. This type of research is quite valuable in that the frontiers of knowledge about that particular court are expanded; however, the limitation of this work is that the ability to apply the findings to other courts may be limited. Nonetheless, in the past several decades, researchers have made significant progress in the field of comparative judicial politics by conducting increasing numbers of both single-court and cross-national judicial studies.

Comparative Judicial Decision Making

A great deal of research in law and courts scholarship deals with judicial decision making. In other words, one of the central endeavors for political scientists who study judges and courts is to explain and predict how and why judges rule in certain types of cases. The attempt to understand this particular form of judicial behavior has been limited largely to analyses of appellate courts in the U.S. system. However, an increasing amount of work is being devoted to examining this phenomenon in comparative perspective. Research has been carried out analyzing the high courts of, among others, Latin American, European, and Middle Eastern nations, as well as Australia and Canada. In addition, courts in both democratic and authoritarian regimes have been studied (see, e.g., Ginsburg & Moustafa, 2008). Many political scientists have sought to test the judicial behavior

of appellate court judges outside the United States by using one of the primary theories of judicial decision making: the *legal*, *attitudinal*, and *strategic* models. The sections that follow provide a brief overview of each of these approaches, as well as the utility of each for explaining judicial behavior across nations and courts.

The Legal Model of Judicial Decision Making

The legal model is the traditional theory of how judges in common law legal systems decide cases. In its simplest form, the legal model proposes that judges make decisions on the merits of a particular case by interpreting the facts of the case through referring to the plain meaning of the relevant statute or constitutional provision, precedent in prior cases, and legislative and original intent of those who drafted the law or constitution (Segal & Spaeth, 2002). So the legal model posits that judges will not use their personal political or partisan values when deciding cases, but only legal factors. In other words, according to the theory, judges will weigh and balance prior precedents, statutes, and constitutional provisions when deciding the outcomes of cases. Again, the individual political ideology of a particular judge is not expected to influence the decision on a case, because judges will focus exclusively on the legal factors. Indeed, judges themselves almost always maintain that they decide cases only according to legal factors, and they usually deny allowing personal policy preferences to influence their voting (Spaeth, 1995). Thus, the legal model of judicial decision making holds that the law, not personal political preferences, will drive the decision-making process of judges.

In comparative perspective, the legal model is not well suited to explain judicial behavior outside common law nations. Because judges in civil law legal systems generally do not rely on the precedent established in previous cases, the legal model as usually formulated does not apply in these courts. It is possible that a modified legal model could be formulated to measure the behavior of judges in civil law courts. However, to date, almost all political scientists have abandoned the legal model framework when studying the behavior of judges outside common law legal systems.

The Attitudinal Model of Judicial Decision Making

Perhaps not surprisingly, many American political scientists have rejected the legal model of judicial decision making and instead have embraced an alternative theory: the attitudinal model. The attitudinal model received its most thorough treatment in Jeffrey Segal and Harold Spaeth's 1993 book, *The Supreme Court and the Attitudinal Model*, and their follow-up volume, *The Supreme Court and the Attitudinal Model Revisited* (2002). Segal and Spaeth argue that the attitudinal model represents a melding of

key concepts from legal realism, political science, psychology, and economics (2002).

Stated most simply, the theory holds that judges tend to vote according to their political preferences, attitudes, and ideologies and not according to the relevant doctrine, legislative intent, or legal precedent of a particular case. Spaeth (1995) contends that the justices vote as they do because they want their decisions to reflect their individual policy preferences. A number of factors are present at the U.S. Supreme Court that encourage attitudinal voting by judges, and these factors generally apply to judges in other common law high courts as well.

First, Supreme Court justices have no ambition for higher judicial office, so the judges usually do not have to consider the consequences of their decisions on the merits of a case (Segal & Spaeth, 2002). In other words, because the U.S. Supreme Court is the highest tribunal in the system, the judges do not have to worry about having their record scrutinized and possibly being denied a promotion to a higher court in the future. Nor do U.S. Supreme Court judges typically seek political office, so there is no fear of political accountability either. Segal and Spaeth also note that Supreme Court justices are immune from electoral accountability. Although many U.S. states employ periodic judicial elections, federal judges enjoy life tenure, and thus there is no fear of losing their office in an election because of their voting record.

Segal and Spaeth (2002) also observe that the U.S. Supreme Court is a court of last resort that has complete control of its docket. That is, the members of the Supreme Court have nearly total discretion as to which cases will be accepted for judicial review. This ensures that cases that are completely without merit or are totally one-sided will usually be screened out. Only those cases that present a genuine dispute between two legally defensible positions will be accepted by the Court. So, the fact that the members of the Supreme Court control their caseload usually ensures that only cases with solid legal justification for both litigants will be accepted, which should encourage attitudinal voting. Furthermore, because the Supreme Court is at the top of the judicial pyramid in the United States, there is no possibility of a higher court's reversing the Supreme Court's decision on a particular case.

Because of all these institutional arrangements—life tenure for its judges, control of the Court's caseload, lack of political or electoral sanctions, court of last resort—the judges at the U.S. Supreme Court are completely free to decide cases on the basis of personal attitudes, values, and ideologies, not the law, according to the attitudinal model of judicial decision making. Spaeth (1995) does note that there are some potential limitations to the attitudinal model. First, it is unclear whether the attitudinal model of judicial decision making applies to lower appellate courts, because of the restrictions under which these judges operate. Specifically, the possibility for judicial (or political) promotion could limit the tendency of these judges to vote their true preferences, as could the possibility of reversal by a higher court. Furthermore, most lower appellate courts do

not have complete control over their docket, which means that many cases lacking a genuine legal conflict must be accepted in these courts.

Whether the attitudinal model can be used extensively to analyze judicial behavior outside the United States remains an open question. Empirical research on high courts in Canada and Australia has demonstrated that the attitudinal model can be used to analyze judicial behavior in those nations and that attitudinal voting has been empirically observed in courts outside the United States (see, e.g., Ostberg & Wetstein, 2007; Weiden, 2010). However, the attitudinal model may be difficult to apply in other nations' courts because of the difficulty in discerning the ideology of individual judges in these systems, and also because some high courts do not release individual-level opinions but rather rely on one single court-level judgment. In other words, if a high court does not indicate in its opinions which judges voted with the majority and which judges dissented, then it may be difficult to analyze individual judges' behavior accurately, which is a necessary component of the analysis of judicial decision making with the attitudinal model.

The Strategic Model of Judicial Decision Making

The third theory of judicial behavior is the strategic model. The strategic model borrows some of its central assumptions from the field of economics and, more specifically, from what is known as *rational choice theory*. Stated simply, the strategic model accepts that judges will vote according to their policy preferences, but in order to maximize those preferences, judges will take into account the voting of the other members of the court and thus cast their votes strategically (see generally Epstein & Knight, 1998). In other words, strategic judges are not unconstrained actors but must weigh the preferences of other actors when making decisions on cases. Many attempts to formally model strategic decision making at the U.S. Supreme Court rely on game-theoretical accounts to demonstrate that some element of judicial decision making was "sophisticated" (i.e., took into account the preferences of other judges) rather than "sincere" (i.e., was a pure example of a judge's voting for his or her policy preferences, without regard to other judges' actions).

There is continued debate over the utility and explanatory power of the strategic model and rational choice theory overall (see generally Green & Shapiro, 1996). However, some of the strongest research in comparative judicial politics has used formal models to analyze and explain the behavior of judges outside the United States (see, e.g., Helmke, 2002; Vanberg, 2001). Because the strategic model of judicial behavior does not rely exclusively on individual-judge-level data that may not always be available, it is well suited to the analysis of comparative judicial politics. Indeed, these formal models of judicial behavior are crucial in generating theories of the judicial process that can be tested empirically against actual court data.

Constitutional and Appellate Courts in Comparative Perspective

Recent work on constitutional courts outside the United States has demonstrated that some political systems are becoming increasingly “judicialized” (Stone, 1992a, 1992b; Tate & Vallinder, 1995). That is, legislators and executives must anticipate the rulings of high courts and frequently alter their policies in advance of adoption to avoid nullification (Tate & Vallinder, 1995). Of course, this phenomenon has always existed to some degree, but it seems clear that this tendency has increased to the point that it can be said with confidence that, in many nations, the political process has been “constitutionalized” (Shapiro & Stone, 1994). Indeed, some commentators have noted that European constitutional courts may now operate as third legislative chambers (Shapiro & Stone, 1994; Stone, 1992b). In some parliaments in which the representatives are weak and dominated by the executive, the constitutional court may be a more effective policy-maker than the parliament itself, in this new era of the judicialization of politics (Stone, 1992a). Martin Shapiro and Alec Stone (1994) note that judges in some European constitutional courts “actually provide the draft statutory language that the judges say they would find constitutional” (p. 404) after reviewing a particular policy. Although the policy-making authority of constitutional courts varies by country, the fact that these courts are well established as policymakers is undeniable.

The judicialization of politics has had other consequences, as well. Scholars have observed that a constitutional politics of rights has become increasingly present, especially in Europe. Shapiro and Stone (1994) describe the constitutionalization of politics as follows:

[It] easily comes to infect the entire political system because opposition political parties, lawyers, citizen groups, and others can see that rights claims are an effective avenue of social change. These actors have become, in essence, the political constituencies of the judges and of constitutional review. (p. 417)

The increase in rights claims and rights discourse has been observed outside Europe as well. Charles Epp (1998) argues that the explosion in rights-based litigation is due primarily to what he terms the “support structure” for legal mobilization, rather than bills of rights, judicial independence, judicial leadership, or any structural explanations. By “support structure,” Epp refers to the financial resources and legal expertise that allow litigants to pursue claims that they almost certainly could not finance on their own. Building on the well-developed legal mobilization literature in the United States, Epp adds another piece to the study of comparative courts.

No discussion of comparative constitutional politics would be complete without mention of the supranational development of constitutional politics. The case law of European Union tribunals such as the European Court of Justice, the European Court of Human Rights, and the

International Court of Justice is often used as precedent in cases before European constitutional courts. Additionally, the European Conference of Constitutional Tribunals, which is composed of all the presidents of the highest courts in Europe, meets regularly at conferences in which their rulings are discussed among each other. Thus, the constitutionalization of the political process is proceeding within national boundaries and is also being encouraged at the supranational level. The tendencies have combined to increase the judicialization of politics in Europe and throughout the world.

Judicial Review in Comparative Perspective

Judicial review is the ability of a court to overturn a piece of legislation, an act of the executive, or a lower court decision. However, judicial review powers vary widely by country and by court. A number of studies have examined judicial review comparatively in an attempt to provide a better understanding of how the practice of judicial review differs around the globe. Indeed, Donald Jackson and Tate’s (1992) *Comparative Judicial Review and Public Policy* provides a series of useful essays describing and comparing the differing forms of judicial review found in the world. The authors note two important distinctions in types of judicial review:

1. *A posteriori* (or *concrete*) review, in which judicial review occurs only after the statute has already taken effect and there is a concrete case or controversy extant, versus *a priori* (or *abstract*) review, in which judicial review may take place before a law takes effect and thus without an actual case or controversy. Stone (1992a) notes that this dichotomy is not exactly accurate, as abstract review can exist *a posteriori*, as in Austria, Portugal, Spain, or West Germany, or *a priori*, as in France.
2. The *all courts model of judicial review* (used in the United States), in which any court at any level may declare a statute unconstitutional, versus the *constitutional courts model of judicial review*, wherein only a specially designated court may review a challenged statute’s constitutionality (Tate, 1992).

In addition, an excellent cross-national study of the formation of constitutional review in new democracies, demonstrating how judicial review serves the interests of political actors during regime change, is found in Tom Ginsburg (2003). Another theory analyzing cross-national constitutionalization is described by Ran Hirschl (2004). This theory is termed *hegemonic preservation*, and it holds that the growth of judicial power is explained by interplay of the self-interest of political actors, economic elites, and judicial notables—all of whom work to create constitutional reform in a manner that serves their own agenda.

Another trend in judicial review in comparative perspective deals with the introduction of the international European Convention on Human Rights. This international agreement may be having a significant impact on the jurisprudence of the United Kingdom, which has traditionally eschewed judicial review of legislation in favor of the centuries-old tradition of *parliamentary supremacy*.

David Weiden (2009b) empirically analyzes recent court cases in Northern Ireland and finds that litigants are now willing to invoke the protections of the European Convention on Human Rights in domestic courts and request that British laws be overturned through the process of a *declaration of incompatibility*. In other words, the long-held principle of parliamentary supremacy in Britain may be slowly giving way to the principle of judicial review, which has never truly existed in the United Kingdom. To be sure, this is a very recent development in Britain, and additional research in coming decades and in other British courts will allow for further confirmation of this trend.

Judicial Independence in Comparative Perspective

The concept of judicial independence is closely related to judicial review and has been another area wherein a great deal of comparative research has been conducted. *Judicial independence*, in its simplest sense, refers to the degree of freedom that judges have from other political actors. In other words, the study of judicial independence examines whether judges can make decisions without being influenced by policymakers and elected officials. A number of institutional features can increase judicial independence. For example, life tenure for judges (rather than fixed terms) ensures that they can decide cases without fear of retaliation from the executive or legislature. However, the ability for judges to make decisions freely is a double-edged sword because judicial decisions that diverge too greatly from public opinion can result in a loss of support for the court, perhaps damaging the court's legitimacy. This danger is especially evident for judges in emerging democracies. Ginsburg (2003; see also Epstein, Knight, & Shvetsova, 2001) found that, in the early years of a court's existence, judges in these developing nations are less likely to make rulings that challenge other political actors.

Party Capability Research

A continuing debate in law and courts scholarship has been whether litigants with more resources are more likely to win their cases than are those who do not possess such resources, an issue referred to as *party capability*. This debate was first described in Marc Galanter's (1974) classic article that examined this question in American appellate courts and found that litigants that are "repeat players" (as opposed to "one shotters") and have greater resources are more likely to prevail in appellate litigation.

The question of whether parties with greater resources tend to prevail in litigation is well suited to comparative judicial analysis. Peter McCormick (1993) examined whether governmental litigants are more successful than private litigants in litigation and found that repeat players are more likely than one shotters are to experience success in the Supreme Court of Canada. However, in Australia, a similar study found that there is little evidence to support the conclusion that litigants with greater resources are

more successful at the High Court (Smyth, 2000). The difference in the results between the Canadian and Australian high courts regarding party capability is striking and raises the question of whether the party capability thesis applies only to certain courts.

The most comprehensive comparative examination of Galanter's (1974) thesis to date is by Stacia Haynie, Tate, Reginald Sheehan, and Donald Songer (2001). These authors examined party capability theory in the high courts of Australia, Canada, England, India, the Philippines, South Africa, and Tanzania. They found that the party capability thesis was strongly supported in the high courts of Canada, Great Britain, and Tanzania but less so in the other cases in the study. Thus, the question as to whether the party capability thesis is generalizable across nations and courts remains open. Further research in additional countries and judicial systems may help answer these questions and constitutes a promising area for additional comparative courts research.

Trial Courts and Juries in Comparative Perspective

In contrast with the well-developed literature on appellate courts, political scientists have conducted little research on trial courts in comparative perspective. Indeed, most of the empirical and interpretive work on comparative trial courts has been done by economists, criminal justice specialists, and law professors. This is somewhat surprising because there are large differences in the trial procedures and processes used in Anglo-American countries and those used in the rest of the world, and these differences present many interesting empirical questions. The trial system used in Anglo-American nations is known as the *adversarial model* and is premised on the assumption that the truth in a court case will be revealed most efficiently through a competition between the litigants, and then the ultimate decision in the case will usually be made by a jury. By contrast, the trial system generally used in Continental Europe, Africa, Latin America, and Asia is the *inquisitorial model*. This trial system is less confrontational than the adversarial model, but instead is more akin to an investigation wherein the judges control the proceedings, call and question the witnesses, and also make the ultimate determination in the case.

Clearly, the fundamentally different approaches toward the administration of a trial found across the globe present intriguing comparative issues for social scientists. The economists Bruno Deffains and Dominique Demougin (2008; see also Block & Parker, 2004) conducted experimental research analyzing whether the adversarial or inquisitorial trial system tended to be more equitable; they found that the adversarial trial system could lead to inequality and inefficiency, at least in criminal cases. These researchers naturally brought an economic focus to their studies; political scientists should be able to add additional political and legal variables to these analyses.

Another area that has been little studied is that of juries in comparative perspective. The trial jury was developed in

England and is still found there, as well as in former British colonies: the United States, Canada, Australia, New Zealand, Ireland, and dozens more nations around the world. Neil Vidmar (2000) reports that the trial jury has been eliminated in most systems for civil trials but is still widely used in criminal cases in those nations retaining a jury system. Given that the trial jury remains ubiquitous in most common law legal systems, a comparative analysis of juries could yield numerous insights for comparative judicial researchers. However, to date, the comparative study of juries remains largely the domain of law professors. Vidmar's *World Jury Systems* is the single best volume on comparative jury research, consisting of a number of essays describing the jury system in particular countries, as well as an invaluable introductory essay. However, true empirical cross-national research on juries remains to be conducted.

Future Directions

The field of comparative judicial systems is still relatively young, and many avenues remain available for new empirical and interpretive research. As noted above, cross-national empirical judicial research (as opposed to single-country studies) remains less common within the field of political science (but see, e.g., Epp, 1998; Herron & Randazzo, 2003; Smithey & Ishiyama, 2002; Weiden, 2010). However, it is likely that the release of the *High Courts Judicial Database* (Haynie, Sheehan, Songer, & Tate, 2007a) will stimulate a great deal of cross-national courts scholarship. This new database consists of numerically coded information for all cases (or a random sample of cases in India and Philippines) reported by the high court in 11 countries: Australia, Canada, India, Namibia, the Philippines, South Africa, Tanzania, the United Kingdom, the United States, Zambia, and Zimbabwe. The period for which the data were collected varies by country, ranging from 9 to 52 years. In other words, each case formally reported by the high court in those nations for the specified period was numerically coded for a number of variables, thus permitting statistical analysis of long-term trends and other phenomena (Haynie, Sheehan, Songer, & Tate, 2007b). One concrete example of this is the emerging research in comparative party capability (Haynie et al., 2001), which was made possible by the development of the *High Courts Judicial Database*.

The ready availability of this data allows researchers to compare various phenomena, such as rates of judicial activism or judicial decision making, in cross-national perspective. The growth of these lines of research may allow for the development of genuine cross-national theories of the judicial process. In other words, it may be possible for scholars to identify certain patterns of judicial behavior that transcend individual court systems. In time, there may be general theories of common law and civil law courts that have been empirically analyzed and confirmed. The role of courts in authoritarian regimes is also an area in which significant progress may be made (see, e.g., Moustafa, 2007).

Another promising direction for future research is the comparative analysis of trial courts and trial systems. As noted above, little research has been conducted by political scientists regarding the differences between the adversarial and inquisitorial trial systems. This fundamental divide has long escaped notice by political scientists, most likely because of the difficulty in obtaining data. Obviously, observing trials outside the United States or other English-speaking nations and collecting data from these tribunals can present considerable language and access difficulties. Nonetheless, experimental research simulating adversarial and inquisitorial trial systems may present one solution to this difficulty (see, generally, Weiden, 2009a).

Similarly, another avenue for future research may involve the analysis of comparative juries: whether certain jury systems are more likely to be more efficient or lead to certain outcomes. This may also be an area in which experimental research may prove to be more practical. Related to this, further research could be conducted to examine the role and influence of lay judges in those inquisitorial trial systems that do not use juries. Overall, the examination of trial courts, trial juries, and lay judges may prove to be an area in which considerable progress may be made.

Finally, the continuation of research regarding judicial independence and separation of powers, using primarily formal models, represents an exceptionally promising area and one that should continue to lead to additional generalized theories of judicial behavior. See generally José Maria Maravall and Adam Przeworski (2003). For example, it may be possible to generate a formal theory of the role of judges in presidential and parliamentary systems and then test those models empirically, using the newly available data in the *High Courts Judicial Database* (Haynie et al., 2007a).

Overall, the diversity of approaches used by comparative courts scholars, combined with the increasing availability of comparative data and the willingness of scholars to look beyond the borders of the United States, ensures that this subfield will continue to yield new insights into comparative judicial behavior and the judicial process.

Conclusion

As discussed, the subfield of comparative judicial politics is both young and old. The comparative study of courts was one of the earliest subjects of political science research before fading into relative obscurity. However, the rediscovery of comparative judicial research in recent decades has produced a body of work containing numerous insights into the role of courts in varying types of political systems. This new research has provided important findings in the areas of judicial decision making, separation of powers and judicial independence, judicial review, and party capability. At the same time, much work is left to be done in, among other areas, comparative trial courts and procedures, juries, judicial independence, and the role of courts in emerging democracies. However, the vitality and rigor demonstrated in the new wave of comparative courts research bodes very well for the future.

References and Further Readings

- Becker, T. L. (1970). *Comparative judicial politics: The political functionings of courts*. Chicago: Rand McNally.
- Block, M. K., & Parker, J. S. (2004). Decision making in the absence of successful fact finding: Theory and experimental evidence on adversarial versus inquisitorial systems of adjudication. *International Review of Law and Economics*, 24, 89–105.
- Deffains, B., & Demougin, D. (2008). The inquisitorial and the adversarial procedure in a criminal court setting. *Journal of Institutional and Theoretical Economics*, 164, 31–43.
- Epp, C. R. (1998). *The rights revolution: Lawyers, activists and supreme courts in comparative perspective*. Chicago: University of Chicago Press.
- Epstein, L., & Knight, J. (1998). *The choices justices make*. Washington, DC: CQ Press.
- Epstein, L., Knight, J., & Shvetsova, O. (2001). The role of constitutional courts in the establishment and maintenance of democratic systems of government. *Law & Society Review*, 35, 117–164.
- Galanter, M. (1974). Why the “haves” come out ahead: Speculations on the limits of social change. *Law & Society Review*, 9, 95–160.
- Ginsburg, T. (2003). *Judicial review in new democracies: Constitutional courts in Asian cases*. New York: Cambridge University Press.
- Ginsburg, T., & Moustafa, T. (Eds.). (2008). *Rule by law: The politics of courts in authoritarian regimes*. Cambridge, UK: Cambridge University Press.
- Green, D., & Shapiro, I. (1996). *Pathologies of rational choice theory: A critique of applications in political science*. New Haven, CT: Yale University Press.
- Haynie, S. L., Sheehan, R. S., Songer, D., & Tate, C. N. (2007a). *High courts judicial database*. Retrieved May 5, 2010, from <http://sitemason.vanderbilt.edu/page/exnake>
- Haynie, S. L., Sheehan, R. S., Songer, D., & Tate, C. N. (2007b). *High courts judicial database codebook*, version 1.2. Retrieved June 20, 2009, from <http://sitemason.vanderbilt.edu/page/exnake>
- Haynie, S. L., Tate, C. N., Sheehan, R. S., & Songer, D. R. (2001). *Winners and losers: A comparative analysis of appellate courts and litigation outcomes*. Paper presented at the annual meeting of the American Political Science Association, San Francisco, CA.
- Helmke, G. (2002). The logic of strategic defection: Court executive relations in Argentina under dictatorship and democracy. *American Political Science Review*, 96, 291–304.
- Herron, E. S., & Randazzo, K. A. (2003). The relationship between independence and judicial review in post Communist courts. *Journal of Politics*, 65, 422–438.
- Hirschl, R. (2004). *Towards juristocracy: The origins and consequences of the new constitutionalism*. Cambridge, MA: Harvard University Press.
- Jackson, D. W., & Tate, C. N. (Eds.). (1992). *Judicial review and public policy in comparative perspective*. New York: Greenwood Press.
- Maravall, J. M., & Przeworski, A. (2003). *Democracy and the rule of law*. Cambridge, UK: Cambridge University Press.
- McCormick, P. (1993). Party capability theory and appellate success in the supreme court of Canada, 1949–1992. *Canadian Journal of Political Science*, 26, 523–540.
- Merryman, J. H. (1985). *The civil law tradition: An introduction to the legal systems of western Europe and Latin America* (2nd ed.). Stanford, CA: Stanford University Press.
- Moustafa, T. (2007). *The struggle for constitutional power: Law, politics, and economic development in Egypt*. Cambridge, UK: Cambridge University Press.
- Ostberg, C. L., & Wetstein, M. E. (2007). *Attitudinal decision making in the supreme court of Canada*. Vancouver, Canada: UBC Press.
- Pizzi, W. T. (1999). *Trials without truth*. New York: New York University Press.
- Segal, J. A., & Spaeth, H. J. (1993). *The supreme court and the attitudinal model*. Cambridge, UK: Cambridge University Press.
- Segal, J. A., & Spaeth, H. J. (2002). *The supreme court and the attitudinal model revisited*. Cambridge, UK: Cambridge University Press.
- Shapiro, M. (1981). *Courts: A comparative and political analysis*. Chicago: University of Chicago Press.
- Shapiro, M., & Stone, A. (Eds.). (1994). The new constitutional politics of Europe [Special issue]. *Comparative Political Studies*, 26.
- Smithey, S. I., & Ishiyama, J. (2002). Judicial activism in post communist politics. *Law & Society Review*, 36, 719–742.
- Smyth, R. (2000). The “haves” and the “have nots”: An empirical study of the rational actor and party capability hypotheses in the High Court 1948–1999. *Australian Journal of Political Science*, 35(2), 255–274.
- Spaeth, H. J. (1995). The attitudinal model. In L. Epstein (Ed.), *Contemplating courts* (pp. 296–314). Washington, DC: CQ Press.
- Stone, A. (1992a). Abstract constitutional review and policy making in western Europe. In D. W. Jackson & C. N. Tate (Eds.), *Comparative judicial review and public policy* (pp. 41–58). New York: Greenwood Press.
- Stone, A. (1992b). *The birth of judicial politics in France: The constitutional council in comparative perspective*. New York: Oxford University Press.
- Tate, C. N. (1992). Comparative judicial review and public policy: Concepts and overview. In D. W. Jackson & C. N. Tate (Eds.), *Comparative judicial review and public policy* (pp. 3–14). New York: Greenwood Press.
- Tate, C. N. (2002a). Past, present, and future with the “comparative advantage”: Part I, 1892–1989. *Law and Courts: The Newsletter of the Law and Courts Section of the American Political Science Association*, 12(2), 3–10.
- Tate, C. N. (2002b). Past, present, and future with the “comparative advantage”: Part II, 1990–2002. *Law and Courts: The Newsletter of the Law and Courts Section of the American Political Science Association*, 12(3), 3–13.
- Tate, C. N., & Vallinder, T. (Eds.). (1995). *The global expansion of judicial power*. New York: New York University Press.
- Vanberg, G. (2001). Legislative judicial relations: A game theoretic approach to constitutional review. *American Journal of Political Science*, 45, 346–361.
- Vidmar, N. (2000). *World jury systems*. Oxford, UK: Oxford University Press.
- Weiden, D. L. (2009a). Comparing judicial institutions: Using an inquisitorial trial simulation to facilitate student understanding of international legal traditions. *PS: Political Science & Politics*, 42, 759–763.
- Weiden, D. L. (2009b). Examining the influence of the European Convention on Human Rights in Northern Ireland courts. *Social Science Journal*, 46, 150–163.
- Weiden, D. L. (2010). Judicial politicization, ideology and activism at the high courts of the United States, Canada and Australia. *Political Research Quarterly*. doi:10.1177/1065912909352775.

23

CIVIL SOCIETY

RYAN SALZMAN

University of North Texas

The tendency of individuals to group themselves according to race, geographic location, and interest has been seen as natural by both historical theorists and contemporary political scientists. It is not surprising that groups are viewed in this way by individuals who think about or study politics. Politics often requires a basic recognition of the necessity of groups for political organization either as a single group of citizens, as factions of competing interests, or as separate races that share a common leader. Therefore, it appears nearly impossible to consider politics without considering the effects of groups on the system as a whole.

The interaction of groups and political actors in society is best conceived through the discussion of civil society, social networks, and social capital. In this chapter, each of these terms describes an aspect of a single idea that associations shape social and political life. *Social networks* are often informal organizations of individuals that span diverse segments of society (Gibson, 2001). These networks can be small or large, but their ultimate purpose becomes promoting the common interest of the network. Within these networks, social capital can be accumulated. *Social capital* is “the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition” (Bourdieu, 1985, p. 248). The accumulation of social capital encourages individuals to act together to achieve common goals.

Without social capital, the achievement of those goals would be impossible. The ebb and flow of social capital accumulation, especially within the context of social networks, creates social associations that interact at various levels of society and government with varying degrees of formality. This broad condition has come to be understood as *civil society*. Social networks interact with other social networks within the purview of civil society. The actions of individuals within civil society promote increases and decreases in social capital that affect future interactions of individuals and social networks. While these terms are separate in what they specifically represent, the central theme remains consistent that relationships matter.

Although historically inseparable from politics in general, the formal discussion and exploration of groups in society have developed with the spread of democracy since the early 19th century. With democracy came the need to consider the preferences of individuals. Often, these preferences were recognized and shared by individuals, who thus organized themselves into groups based on those shared preferences. The desire of political leaders to know and respond to these groups has promoted the study of associations in society.

This chapter looks at the theoretical and empirical pursuits of research involving the study of social capital, social networks, and civil society. Before delving into the contemporary ideas that are the primary subject of this chapter, it is important to discuss the role of associations

and groups in political life as it has been seen historically. Next, the theoretical concepts of social capital, social networks, and civil society are laid out separately, with attention given to how the terms are interrelated. Then, the relationship of the concepts to the study of democracy is examined as the primary line of research involving social capital, social networks, and civil society. The next section focuses on social capital and civil society as they relate to the economy and the broader society. Then criticisms and avenues for future research are discussed.

The Importance of Associations in Political Life

From the advent of political society discussed by the ancient Greek philosophers to the earliest observations of American democracy and through the global adoption of democracy in the mid-20th century, the recognition of the centrality of humans associating with each other has been revealed and accepted in the context of participatory politics. For Aristotle (circa 335 BCE), in order for individuals to make the correct decision in choosing leaders, citizens had to know about each other. Lacking this knowledge, it was impossible to make proper political decisions for the community. In this way, Aristotle viewed the city as a group in which individuals interacted to gain knowledge of each other's character and preferences. This interaction was necessary when politics required the participation of citizens.

In his observations of early American democracy, Alexis de Tocqueville (1840) wrote at length on the prevalence and necessity of associations in the young country. He viewed these associations as important to the type of participatory society that had blossomed. In America, associations were engaged more successfully than in any other place in the world at that time. By being involved in associations, American citizens were able to overcome their lack of influence as separate individuals. Tocqueville observed that when individuals with a common opinion met, they naturally combined themselves into an association. As the association grew, political actors were forced to take notice of the association and recognize the preferences of the group members. In this way, associations empowered individuals in the political context, which forced accommodation by political actors. It is these associations that maintain the core of the civil society, social networks, and social capital discussion.

Following the observations of Tocqueville, Émile Durkheim (1893/1984) explored the interactions of individuals within society and observed that connections between individuals remain after the initial interaction. These remaining social ties contribute to the functioning of the community in a manner that is broader than the initial interaction by shaping the condition of social capital that results from the interaction. In this way, associations

persistently affect one another through the lasting impact that individuals make on each other.

Many early political scientists promoted the importance of voluntary associations. For example, Almond and Verba (1963) promoted voluntary associations as the most important mediating factor between individuals and the state. Associating with other individuals in a voluntary association gives a person increased political resources that can be used to achieve his or her desired political ends. Also, membership in associations affects an individual's political attitudes. For Almond and Verba, voluntary associations were as important as they were variable. It is this accepted variation between groups that shaped the inquiries related to civil society, social networks, and social capital in the decades that followed Almond and Verba's research.

A Theory of Social Capital, Social Networks, and Civil Society

Social capital, social networks, and civil society are terms and ideas that are related but not necessarily synonymous. Social capital can be the measure by which social networks and civil society are evaluated, but ultimately, it describes relations between individuals or entities. Civil society is an aggregate perspective of relative social capital per sector, society, or state. Social networks are the building blocks of civil society but are generally considered both autonomous from each other and yet integrated into a common civil society. To understand them all, we must understand each.

Social Capital

While decades of work focused on the role of associations in daily life, the concept of social capital did not become cemented in its contemporary form until Coleman (1988) examined different forms of capital and pronounced that, like other forms of capital, social capital existed and facilitated interactions between individuals or organizations. In comparison with physical and human capital, social capital was seen as the least tangible. Instead of focusing on production or relative skill sets, social capital focuses on the functions of certain aspects of the social structure. "The function identified by the concept of 'social capital' is the value of these aspects of social structure to actors as resources that they can use to achieve their interests" (Coleman, 1988, p. S101). Identifying the functions of the social structure became one way to account for the differences of outcomes for individuals, without having to elaborate on the social structure details through which the transitions occur.

To briefly summarize an example provided by Coleman, clandestine organizations can help promote revolutionary activities where otherwise peaceful protests are the expected outcome. In this scenario, the organization of individuals is the social capital that causes the change to revolutionary output. The system remains unchanged. Instead, it is the

introduction of a form of social capital different from what had already existed that prompts the new output. A stable system can produce highly varied outcomes, depending on the social capital that is invested, along with the systematic resources. Coleman's basic assertion is that producing and integrating social capital to produce certain outcomes is no different from combining a raw material (e.g., petroleum) with various physical capital (e.g., technologies) to produce different products (e.g., motor oil vs. gasoline).

For Coleman, the reality of social capital in social structures is dictated by obligations, expectations, and trustworthiness. Ultimately, this dynamic can be easily defined as a system of general reciprocity in which doing something for an individual will prompt a comparable response by the receiving individual. In this situation, trust in the individual who is being helped creates an expectation in that individual to reciprocate. For the individual being helped, an obligation to return the favor is felt and carried through. It is in this manner that social capital is created and developed. An individual can have a store of social capital that he or she can redeem to shape the outcome of a situation. Trust is the crux of the social capital dynamic. Without trust, social capital is virtually impossible to produce.

For Uslaner (2002), trust is a central consideration in many aspects of human activity. While he accepts that trust is not the only way to achieve cooperation, he embraces the idea that a system of reciprocity is more reliable with higher levels of trust. For social capital to operate in the manner perceived by Coleman, generalized reciprocity must be anticipated. Therefore, trust must be achieved, but it must also be assumed. For other researchers, changes in the dynamic of trust reshape the reality of social capital and thus shape civil society in general.

In an attempt to clarify the role of social capital in affecting governance and the reverse of government affecting social capital, Putnam and Goss (2003) differentiated social capital so that the variation between systems could be more easily judged. Some social capital is formal whereas other social capital is informal. The type of organizational structure (e.g., union vs. supper club) dictates whether social capital is formal or not. Another distinction is between thick and thin social capital. *Thick* social capital exists in tightly connected groups, with *thin* social capital being more prevalent among acquaintances. *Inward looking* social capital focuses on the well-being of group members, whereas *outward looking* social capital focuses on public goods. Finally, *bridging* social capital brings dissimilar individuals together, as opposed to *bonding* social capital, which brings similar individuals together. Although Putnam and Goss's distinctions do not greatly add to the theoretical conceptualization laid out by Coleman, they help to clarify the variations in social capital.

Social Networks

In his seminal works examining civil society and social networks, Putnam adopts the idea of social capital discussed

above. For Putnam, like Coleman, social capital centers on trust and generalized reciprocity. In examining civil societies, he also promotes social capital as a public good that is achievable by everyone in a society. In this way, social capital deviates from other forms of capital, including physical and human capital. Social capital and networks of civic engagement create and maintain norms that can act as natural constraints on individuals' or groups' actions through the recognition and embrace of externalities that result from actions that are either positive or negative for the individuals involved. Therefore, social capital can be created, maintained, and redeemed to affect outcomes in an organized system.

The main avenue through which social capital operates is networks of civic engagement, or social networks. In his study of the development of Italy's economic and social institutions, Putnam (1993) considers the form of networks that existed hundreds of years ago and their effect on institutional outcomes in contemporary Italy. Unlike civil society, which is much broader, social networks are particular to the shared interests of the participating individuals. Individuals can be involved in multiple networks that often overlap. These networks are seen by Putnam as instruments for organizing social capital, with many positive outcomes. In fact, Putnam considers these networks of civic engagement to be social capital themselves. In a contemporary sense, these networks could be parent-teacher associations, service fraternities, political parties, tennis clubs, and the like.

The benefits of social networks are often dictated by the density of the networks. *Density* means the quantity and intensity of individuals and interactions in the group. Basically, increased cooperation for mutual benefit is easier in denser networks and more difficult in less dense networks. Putnam (1993) lists a number of other ways that these networks benefit their members. The defection of individuals from the desired outcome of the network is discouraged through increasing the potential cost of defection for the individual. This encourages cooperation. Robust norms of reciprocity are also encouraged by networks of civic engagement. The degree of trustworthiness of individuals is easier to identify because of the increased levels of communication and information flows that are facilitated by the network. Also, the success of networks in achieving collaboration can serve as a template for future collaboration.

Structurally, networks of civic engagement are horizontal in nature. Horizontal networks are not structured hierarchically but are instead dispersed across society in a manner that renders positions in the various networks as roughly equivalent in terms of power. This is not to say that they are always structured in this way, but they are best at achieving the above stated goals when they are horizontal. According to Putnam, only through horizontal networks can social trust and cooperation be achieved. Properly organized social networks can encourage social capital that can improve the efficiency

of society through the encouragement of coordinating actions.

Civil Society

In his writings on the development of Italian society and the changes in the American context, Putnam (1993, 1995, 2000) is most interested in examining the trends of social capital in civil society. For civil society, the focus is on the community in general. Unlike the direct relationships between individuals and entities that are most important for social capital, or the interest-specific networks of civic engagement, civil society is a broad picture of the community. When one is studying civil society, it is important to understand that degrees of connectedness between individuals and entities and within social networks can have society-wide effects that can result from aggregate levels of civic engagement.

Like the positive effects of social networks that were discussed above, the positive externalities of associations and social capital accumulation can affect even those individuals with zero social capital, or those who are involved in no networks. Communities that have generally higher levels of connectedness and social capital stock often find collective action to be easier. This is due to the pervasiveness of the norms that are developed between individuals and within networks. Therefore, trends that affect individuals will subsequently affect entire networks or sectors and ultimately the entire society.

This idea of a broad civil society is addressed in Putnam's 1995 journal article "Bowling Alone: America's Declining Social Capital." In his research, Putnam points to the reality of increased political disengagement as having originated with changes in individual-to-individual interaction, as well as broader changes in the construct of social networks. These changes, he contends, subsequently reshaped the orientation of America's civil society.

Putnam derives his research question (why has civil society changed?) from a seemingly unrelated observation of political participation. He notes that political participation, particularly voting, has steadily declined in the United States. In trying to understand why this has happened, he looks to changes that have occurred in civil society. From this examination, he draws a simple observation on which he rests his argument: U.S. citizens are bowling now more than ever, but participation in bowling leagues is lower now than ever before. Like the examples of social networks, or networks of civic engagement, discussed above, bowling leagues can promote the creation and accumulation of social capital. Putnam moves from this specific observation to a broader set of observations, all reaching the same conclusion that social networks are changing and declining. Reverting back to his understanding of social capital, he contends that a negative shift in "neighborliness" and "social trust" has accompanied the decline in civic engagement. While the direction of the causal arrow is left to future research, Putnam uses this reality to describe the

aggregate-level situation that is present in U.S. civil society. Although he gives a number of potential explanations for this change, what is certain is that it has happened.

As the above discussion of the theories surrounding the concepts of social capital, social networks, and civil society makes clear, each of the concepts is related to the others and yet carries its own distinct characteristics for whom it pertains to and what functions it affects. Social capital is a building block composed of trust and generalized reciprocity through which relationships are established and maintained. Social networks expand the relationships so that they exist between more than just two individuals and are naturally grown by the light of common interests. The extent to which social capital is embraced through social networks shapes the general understanding of a system's civil society. The degree of interconnectedness in a system affects the relative development of that system's civil society, which in turn affects many other aggregate conditions of society at large.

The next section looks at the main line of research in which civil society and social capital are considered: democracy. From this research, tangent lines of research have also developed to consider the impact of social capital, social networks, and civil society. To round out the discussion, criticisms and future research areas are explored.

Civil Society, Social Capital, Social Networks, and Democracy

In participatory political systems (i.e., democracy), the context within which individuals operate dictates their perspective and subsequent action in that system. Just as individuals with larger amounts of wealth are expected to act differently from individuals with meager amounts of wealth, so too are individuals in groups expected to act differently from individuals not in groups. In this intuitive reasoning, it is easy to see that associations bring change to the operations of individuals in politics. This is evident in the debate that surrounded the creation of the U.S. political system, in which the federalists and antifederalists supported competing perspectives on how to create the best government to represent the people.

For James Madison, factions were to be guarded against in the new U.S. democracy. In Federalist No. 10, Madison (1787/1982) made clear that it is natural for individuals to organize themselves into associations with others with whom they share similar interests. Outside of suspending the liberty of citizens, the only way to control the effects of factious associations was by creating a republican form of government. This, he imagined, would ameliorate the effects of the factions without suspending liberty or undermining the participatory political system that the U.S. forefathers intended to create. This early example makes evident the profound nature of associations in democracy. For studies looking at the impact of social capital, social networks, and civil society, their effects on government type, and especially democracy, have been the most

central in this line of research. This section considers some of the research that has been conducted looking at the impact of civil society (and its parts) on democracy.

As discussed above, Putnam (1995) observed the decline in voting in the United States and desired to explain why this had happened. He theorized that declines in participation in associations had caused a decline in social capital, which subsequently led to the decline in voting. For Putnam, the decline in voting ultimately stemmed from the disconnect that had been growing between individuals. Without a need to be connected, there was little need to make group decisions. Although his tests are simple, his conclusions are presented as concrete: Decreases in civil society participation caused the decline in voting.

Writing in 1995, Verba, Schlozman, and Brady considered civic life and its effect on politics in general. They asserted that nonpolitical institutions (including associations and social networks) enhanced citizen activity in politics. This happened for a number of reasons. For instance, participants in civil society were exposed to increased political stimuli that enhanced their willingness and ability to participate in politics. Also, participation in civil society encouraged further participation in civil society so that associational engagement had a multiplicative effect for citizens and society.

Associations also dictate much in democratic societies by setting the political agenda (Cohen & Rogers, 1992). They act as intermediaries between individuals and the state. This action, as an intermediary, can give an association specifically, or civil society broadly, a great deal of power in shaping and coordinating the preferences of participants. Although Cohen and Rogers, like Madison, are concerned with curbing the role of factious associations, they admit that groups and networks can often contribute to democratic governance in a positive manner.

Political participation is particularly important in democratic societies and has thus occupied much of the research on social capital, social networks, civil society, and democracy. In their work looking at participation in the United States, Rosenstone and Hansen (2003) see the voluntary aspect of social involvement as being important in citizen mobilization. First, most groups engage in their own political mobilization. Second, groups expose their members to sympathetic politicians and activists who engage in mobilizing the members. Finally, simply being a member of an organization exposes the members to potential rewards that are jointly sought by other members and can be best achieved through political action. In this way, the degree of social capital, social networks, and civil society can variably affect participation in political life.

For Verba, Nie, and Kim (1978), voluntary associations have an effect on citizen participation that is independent of resources or other factors. These associations can modify the exchange of resources in political activity. Verba et al. anticipated that associations should motivate participation, especially voting, to a greater extent than strictly resources would. However, they also found that resources influence association membership. Presaging this sentiment,

Huntington and Nelson (1976) found that organizational membership is important for political participation. While Verba, Nie, and Kim focus on seven countries that tend to be more developed, Huntington and Nelson are strictly considering countries that are less developed.

Following the lead of Huntington and Nelson (1976), many other researchers have focused on the developing world. This distinction is important because it introduces increased variation in outcomes, which are often more stable across developed countries. The best example of the third world context introducing a new wrinkle to the study of civil society is an article written by Booth and Richard (1998). In their research, they identify a sector of society that they term *uncivil society*. This uncivil society is violent and confrontational and often antidemocratic in its associational condition. Like other variations on civil society, this sector affects governmental action and outputs. Unlike more developed countries, uncivil society is more prevalent in developing regions and specifically Central America.

In countries that have recently transitioned to democracy, social networks have been shown to facilitate that transition. Gibson (2001) explores this concept by looking at post-Communist Russia. Having developed out of the closed Soviet society, social networks with weak ties between individuals occupied the position normally held by formal civil society in helping initiate and develop democracy. Gibson finds that, although informal in nature, social networks provided the political discussion and organization that could evolve into a robust civil society. It becomes easy to see that the social capital–civil society relationship is developmental in nature but can be trusted to emerge from even weak social networks.

Beyond facilitating transitions to democracy, the continuation and consolidation of a democratic system are encouraged by civil society. Although many researchers and theorists have focused on the preeminent role of elites in affecting democratic systems, properly functioning democracies undoubtedly require the input of the people. For this reason, Diamond (1999) sees civil society as irreplaceable for democratic success. In many ways, civil society plays an intermediary role between the private sphere and the state. First, civil society focuses on public ends over private ends. Next, it relates to the state but does not seek to control it. Finally, civil society encompasses pluralism and diversity (Diamond, 1999). Representing the interests of the people, civil society's most important role is its ability to check and limit the power of the state while simultaneously helping to reform it. Therefore, civil society can be roundly viewed as a positive influence on democratic systems.

Other Lines of Research

Besides affecting democratic governance, social capital, social networks, and civil society affect the economic and productive aspects of society. In initially laying out the theoretical underpinnings of social capital, Coleman (1988)

was looking to explain the creation of human capital (e.g., through education). Ultimately, he found that the presence of social capital in the immediate interactions that students had with community members and family, as well as the general social capital present in the community, affected relative drop-out rates among students. Thus, he demonstrated that social capital is an important form of capital, like physical, financial, and human capital, for shaping the economic potential of an individual and the community.

For Putnam's (1993) study of Italy, the development of social capital (and subsequent social networks) was crucial in initiating the economic differences in development for northern and southern Italy. The existence of a vibrant civil society in northern Italy, derived from the mutual assistance established by trade guilds, encouraged an efficient economy that fostered economic development. In the south, vertically organized societies never allowed for cooperation and subsequently stunted economic development. For Putnam, the ideas of reciprocity and trust are a cornerstone for positive development in most aspects of society, including the economy.

These concepts of efficiency in the productive aspects of society are echoed by Dekker and Uslaner (2001) in the sociological discussion of the role of social capital in affecting economic outcomes for communities. For these authors, a number of conditions result from stronger social networks and social capital relations. First, information sharing encourages efficiency by allowing individuals to avoid inefficient means of production. Second, activities are more coordinated in communities with higher social capital. The example that the authors give points to the failures of irrigation systems due to self-interested actions by farmers who diverted water because the opportunity existed. Third, collective decision making encourages a more efficient distribution of public goods. For Dekker and Uslaner, all three of these efficiency-encouraging activities result from higher levels of social capital in a community.

Besides the governmental and economic effects of social capital, social networks, and civil society, a number of strictly communal impacts have been identified. The clearest identification of some of these benefits is made by Putnam and Feldstein (2003). Not only are the benefits of social capital individual specific, but they also extend to the community as a public good. Drawing correlations between the degree of social capital and community conditions, Putnam and Feldstein observe that communities with higher degrees of social capital (conceived of and measured differently) also have lower crime rates, healthier newborns, and lower drop-out rates. These improved conditions illustrate social capital's importance in society beyond the scope of economic efficiency and democratic governance.

Criticisms

Many studies have extended the ideas associated with social capital, social networks, and civil society in ways

that are critical but productive. Most criticisms derive from the idea that these terms are incomplete. Work by Foley and Edwards (1996) is a good example of this extension. Beyond the understanding of civil society that this chapter has discussed (i.e., that civil society promotes further civility in the population that subsequently fosters democratic governance), the researchers embrace the inclusion of an alternative civil society. They conceive traditional civil society, called Civil Society I, and the alternative civil society, called Civil Society II.

Civil Society II is more autonomous from the governmental apparatus and is thus in a better position to oppose it. This argument is especially applicable for tyrannical regimes. In that context, Civil Society II is able to oppose the tyrannical regime and potentially encourage regime change. In a similar manner, Booth and Richard (1998) extend the discussion of civil society to include Civil Society III. This form of civil society is limited to strict efforts of regime replacement. Basically, Civil Society III is revolutionary in nature. While each of these criticisms is more productive than not, it is important to focus on the completeness of the concept in all contexts.

Another prominent criticism of the social capital, social networks, and civil society discussion is that the conclusions surrounding the impact of associations are overstated. Although many researchers contend that social capital directly impacts aspects of individual and societal life, others disagree, saying that it is only a part of the puzzle. For instance, Stolle and Hooghe (2003) question the importance of associational life for individuals. Also, they conclude that the impact of social capital outside of the group setting is weak, if present at all. Although they do not desire to remove associational life from its prominent position of study, they encourage a distinct focus on the internal workings of groups (and the subsequent social capital movement). Also, Stolle and Hooghe push for further study of the external impacts of groups and the role of civil society.

Another criticism that often arises in studies of social capital, social networks, and civil society is that of measurement of the topic in focus. How does someone accurately measure an idea such as social capital? Approaches to measurement abound. Some individuals prefer surveys to identify social capital. Not without problems, surveys can often make cross-country comparisons impossible because of a lack of transferability of ideas. For civil society, many researchers employ the number of groups in a society or the number of groups to which an individual belongs on average. Again, problems of measurement emerge pertaining to what groups to include, how to distinguish between group types, and whether to include a time element in the measure. Social networks also pose difficulties in measurement due to the intrinsically weak nature of the connections between individuals. Although these issues are not insurmountable,

they must be considered in evaluating the conclusions about the concepts that are drawn by researchers.

Future Research

Research on social capital, social networks, and civil society has come a long way since the work of Coleman and Putnam 20 years ago. However, as the concepts have continued to develop, the context of society has continued to change. With this change comes a need to continue developing the concepts and test for changes in their applicability. One example of a new concept that may shed light on the role of social capital in affecting democratic behavior is that of *political capital*. According to Booth and Richard (2009), political capital is the linkage mechanism that connects social capital to political outcomes. Studies like this will help clarify the effect of social capital and civil society on democracies.

One major change that has occurred over the past decade, and that will cause a particular need for further research, is the development of technology, which has reshaped how we perceive groups to be structured. In his study of the United States, Putnam (1995, 2000) points to the technological development of television as a cause of the decline in neighborliness, which directly affected civil society. The advent of what he calls *tertiary associations*, understood as associations that involve mostly mailings and dues payments but few meetings, changed the way people in America viewed groups. In these ways, groups and their impact on politics and society changed over time.

Technological innovations that will again change the impact of groups on politics and society most notably involve the Internet. Although interactions among group members are more direct on the Internet than in the case of tertiary associations, it is uncertain whether social capital developed through the Internet will act similarly to that developed in community associations such as the parent–teacher association at a local school. In many ways, social networking sites are exactly as advertised: social networks. In other ways, these sites appear to encourage individualism before generalized reciprocity. For these reasons, research focusing on the role of social capital, social networks, and civil society must continue.

Without anticipating technological innovations, is there still a place for research as it is being conducted now? The answer is a resounding yes. The more classical perspective on the roles of associations in society still demands a great deal of attention in the developing world. Many countries continue to strive for democratic governance, and many others are pushing for the consolidation of their democratic regimes. In these countries, studies of traditional civil society and social networks still have much fruit to bear. Besides confirming already held perspectives, these studies will allow for new contextual realities to be integrated into past studies, which

will increase the robustness of or inspire revisions to the conclusions reached by researchers. Basically, there is still much to learn about social capital, social networks, and civil society in nondemocracies, new democracies, and consolidating democracies that cannot be learned in already consolidated democracies.

Conclusion

Individuals naturally associate with each other. Beyond simple, inconsequential interactions, people often seek out groups to join based on their preferences. The basis for joining can be religious, ethnic, political, and so forth. Associations can be in the form of environmental groups, bowling leagues, churches, political parties, neighborhood associations, social networking Internet sites, and the like. Groups are as varied as the people who compose their membership. However, even with the degree of variation that exists, there are certain consistent effects for individuals that derive from being a member of a group. These effects are conceived of as social capital. As members of groups, individuals build social capital by contributing to a system of generalized reciprocity in which other individuals, often members of the same group, will then return the effort that the individual contributed. That return may be in a similar context in which it was given. However, that return may also manifest itself in a unique context. Regardless, individuals desire to reciprocate what others have given. In this manner, social capital affects group and individual performance by inspiring trust and a desire to return the favor.

Social networks and civil society are broader concepts that focus more on the types and numbers of groups within a specific system (e.g., a country). Within social networks and civil society, social capital is exchanged and stored so as to affect individuals' attitudes and behavior. Social networks emerge as a less formal organization of individuals with common interests who, when acting together, can have a greater effect on the system. Civil society is the most aggregate concept and describes the number and intensity of social networks and other associations in a system. The strength of civil society is a broad conceptualization of the associational nature of a system that may affect much of the behavior of system-level actors.

The concepts of this chapter are most effective in countries with democratic regimes. In these participatory political systems, associations become another avenue of participation and can inspire or invite mobilization in the political system. Political participation shapes democratic governance, and associations can shape participation. Besides politics, a strong civil society can also affect the economy by encouraging efficiency and affecting general demographic conditions through educational attainment and reducing criminal tendencies of individuals in a community.

The academic endeavors of researchers have spanned the entire globe. While the context of research is highly

variable between developed and developing countries, the general consensus remains that social capital, social networks, and civil society are consistent concepts. The primary variation is the distinction between participatory and nonparticipatory political systems. Basically, systems that encourage individual input (i.e., democracy) cannot avoid influence by associations. That influence can be intense and broad, just as it can be subtle and specific. Only with continued research will those distinctions become obvious.

References and Further Readings

- Almond, G. A., & Verba, S. (1963). *Civic culture*. Princeton, NJ: Princeton University Press.
- Blaney, D. L., & Pasha, M. K. (1993). Civil society and democracy in the third world: Ambiguities and historical possibilities. *Studies in Comparative International Development*, 28(1), 3–24.
- Booth, J. A., & Richard, P. B. (1998). Civil society, political capital, and democratization in Central America. *Journal of Politics*, 60(3), 780–800.
- Booth, J. A., & Richard, P. B. (2009). Descifrando el capital social en las democracias de América Latina [Deciphering social capital in the democracies of Latin America]. In L. H. Gómez & J. A. Booth (Eds.), *Perspectivas para la democracia en América Latina* [Perspectives on democracy in Latin America]. Toluca, Mexico: Universidad Autónoma del Estado de México.
- Bourdieu, P. (1985). The forms of social capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood.
- Cohen, J., & Rogers, J. (1992). Secondary associations and democratic governance. *Politics & Society*, 20(4), 393–472.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94(Suppl.), S95–S120.
- Dekker, P., & Uslaner, E. M. (2001). *Social capital and participation in everyday life*. London: Routledge.
- Diamond, L. (1999). *Developing democracy: Toward consolidation*. Baltimore: Johns Hopkins University Press.
- Durkheim, É. (1984). *The division of labor in society* (W. D. Halls, Trans.). New York: Free Press. (Original work published 1893)
- Edwards, B., Foley, M. W., & Diani, M. (2001). *Beyond Tocqueville: Civil society and the social capital debate in comparative perspective*. Hanover, NH: University Press of New England.
- Field, J. (2003). *Social capital*. London: Routledge.
- Fine, B. (2001). *Social capital versus social theory: Political economy and social science at the turn of the millennium*. New York: Routledge.
- Foley, M. W., & Edwards, B. (1996). The paradox of civil society. *Journal of Democracy*, 7(3), 38–52.
- Gibson, J. L. (2001). Social networks, civil society, and the prospects for consolidating Russia's democratic transition. *American Journal of Political Science*, 45(1), 51–69.
- Grootaert, C. (2001). Social capital: The missing link? In P. Dekker & E. M. Uslaner (Eds.), *Social capital and participation in everyday life* (pp. 9–29). New York: Routledge.
- Huntington, S. P., & Nelson, J. M. (1976). *No easy choice: Political participation in developing countries*. Cambridge, MA: Harvard University Press.
- Isham, J., Kelly, T., & Ramaswamy, S. (2002). *Social capital and economic development: Well being in developing countries*. Northampton, MA: Edward Elgar.
- Madison, J. (1982). Federalist No. 10. In G. Wills (Ed.), *The federalist papers by Alexander Hamilton, James Madison, and John Jay* (pp. 42–49). New York: Bantam Books. (Original work published 1787)
- Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 4(1), 65–78.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- Putnam, R. D., & Feldstein, L. M. (2003). *Better together: Restoring the American community*. New York: Simon & Schuster.
- Putnam, R. D., & Goss, K. A. (2003). Introduction. In R. D. Putnam (Ed.), *Democracies in flux: The evolution of social capital in contemporary society* (pp. 3–20). Oxford, UK: Oxford University Press.
- Rosenstone, S. J., & Hansen, J. M. (2003). *Mobilization, participation, and democracy in America*. New York: Longman.
- Simpson, P. L. P. (1997). *The politics of Aristotle*. Chapel Hill: University of North Carolina Press.
- Stolle, D., & Hooghe, M. (2003). Conclusion: The sources of social capital reconsidered. In M. Hooghe & D. Stolle (Eds.), *Generating social capital: Civil society and institutions in comparative perspective* (pp. 231–248). New York: Palgrave Macmillan.
- Tocqueville, A. de. (1840). *Democracy in America*. New York: Schocken.
- Uslaner, E. M. (2002). *The moral foundations of trust*. New York: Cambridge University Press.
- Verba, S., Nie, N. H., & Kim, J. (1978). *Participation and political equality: A seven nation comparison*. Chicago: University of Chicago Press.
- Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

24

POLITICAL CULTURE

MARC HOOGHE

University of Leuven

The concept of political culture refers to the political attitudes and behavioral patterns of the population, and it is assumed that this culture largely determines the relation of citizens with the political system. Most studies on political culture claim that specific elements of that culture have an impact on the way political institutions function, although it has to be noted that the reverse causal logic (institutions determining the political culture) has been argued as well by authors adhering to an institutionalist perspective on politics. Political culture includes both the individual's view of himself or herself as a competent political actor and the perception about his or her role within the political system. Strictly speaking, *political culture* refers only to the attitudes of citizens, but in practice it also includes behavioral patterns that are closely related to these attitudes. This chapter first reviews the development of studies on political culture, paying specific attention to the work of Gabriel Almond, Sidney Verba, and Robert Putnam. Subsequently it reviews the empirical research on specific elements of political culture before closing with a glance at future directions in this subfield of political science.

Studies on Political Culture

The Civic Culture

Very few subfields in political science have been determined so strongly by just one book as has the study of political

culture. That book is *The Civic Culture* (TCC), by Almond and Verba, which first appeared in 1963. In this book, the authors argue that a specific orientation toward politics is crucial for maintaining the institutional status quo of democratic political systems. In this regard, Almond and Verba argue for a strong culturalist approach to the study of democratic stability. It is assumed in their work that the presence of a political culture is responsible for the effectiveness and stability of a political system. As such, they oppose various forms of institutionalism, arguing that democratic stability is enhanced first of all by the presence of strong and effective institutions and constitutional rules.

Almond and Verba distinguish three different phases in the development of political cultures. A first, more traditional form most often found in closed traditional societies is the *parochial political culture*. This kind of political culture is very strongly locally based and is focused on adherence and deference to a charismatic leader. This leader combines various social roles in that he or she does not exercise just political power; military, religious, and sometimes even medical powers are also attributed to him or her. Second, Almond and Verba list the *subject political culture*. In this form of political culture, roles are already more strongly differentiated, and this culture is compatible with the functioning of nation-states, covering a larger territory than the purely parochial communities. In this kind of system, citizens already have acquired a very distinct role, that of subject, with the duty to obey the commands of the head

of state. This implies that this form of political culture is very closely linked to the development of *absolute monarchy* during the 16th and 17th centuries in western Europe. The monarch holds absolute power and has no need to legitimize that power in public opinion, often claiming historical or even religious sources of legitimacy.

Almond and Verba note that the subject political culture has become unsustainable in the developed democracies of the 20th century. As both the education level and the economic status of Western populations have increased, it has become increasingly difficult to limit the role of citizens to merely obeying the sovereign's orders. A typical modern form of political culture therefore is the *participant political culture*. In this form, citizens assume that they are able to express themselves on political matters, and they also take it for granted that they have the right to participate in processes of political decision making. During the 20th century, the participant political culture grew in power, and it has become almost impossible for the political elite to ignore the demand for more participatory openness.

However, Almond and Verba do not see the participant political culture as the type of culture that is most easily combined with the stability of democratic regimes. Although they acknowledge the fact that political systems should allow for routine participation of citizens, they also express concern about a possible "overload" of the political system. If a large number of citizens want to participate in a routine manner, and if they do not feel inclined to accept the output of the political system in case their demands have not been met, this would mean that the political system in effect can no longer govern. Almond and Verba, therefore, express some concern that citizens should not participate constantly or too intensively. Rather, they are potential participants, who do not interfere constantly but always feel sufficiently efficacious to interfere if the need should arise. Furthermore, they are willing to accept the decisions of the political system, as long as these are the result of the democratic process.

The ideal *civic culture*, therefore, according to Almond and Verba, is a mixture of the three preceding forms. Citizens feel sufficiently empowered to participate in the decision-making process, but they are also loyal to the political system, and they are willing to adhere to the decisions that have been made by that system.

The Almond and Verba study has been hugely successful, and it has continued to dominate the field of political culture studies for decades. There are a number of reasons TCC had such a strong and lasting impact.

First of all, it was one of the first successful applications of comparative survey research. The book included survey material from five different countries, which was very innovative in the early 1960s. As such, Almond and Verba could make more convincing claims about causality and relationships between variables than could earlier studies that focused on just one country. Second, and partly related to this, is that Almond and Verba were the first to be able to

strongly defend the importance of public opinion. Claims about the effect of public opinion self-evidently require access to survey data, and since these data were hardly available before the 1950s, Almond and Verba could make a much more powerful claim than previous authors had done. One might even say that, by themselves, they reinvigorated the culturalist claim in the study of politics. Third, in the early 1960s, there was indeed quite some concern, both in academia and in society, about the stability of democratic political regimes. Quite a few Western countries felt threatened by the rise of Communist rule in Russia and in Central and Eastern Europe. The ongoing process of decolonization, furthermore, led to the question of how one could ensure democratic rule in the newly independent countries of Africa and Asia. TCC provided an instant answer to all these questions: One should try to install a civic culture among the populations of these countries.

In general, it can be observed that TCC endorsed a very moderate view of political culture that resonated apparently quite successfully with the dominant view in the early 1960s. On one hand, Almond and Verba strongly supported the normative ideal of an active citizenry, but they simultaneously stressed the need for citizens' loyalty to the political system. This latter point especially was increasingly criticized during the 1970s, a period when support for an activist role of citizens became increasingly popular. In that period, TCC was also criticized for what was labeled its focus on Anglo-Saxon political cultures (Inglehart, 1988). The emphasis on citizens' loyalty implies a preference for stability and incremental changes in the political system. Revolutionary or disruptive forms of political participation are not really evaluated in a positive manner in the Almond and Verba framework.

Social Capital and Political Culture

Research on the development of political cultures in modern democracies was strongly reinvigorated by the publication of the 1993 volume by Putnam on civic traditions in modern Italy. To some extent, one could argue that Putnam (1993) essentially took a new look at the question Almond and Verba had posed three decades earlier: What makes democracy work? The innovative feature of Putnam's research was that he could start from a quasi-experimental research setting. In 1970, 20 regional governments in Italy received a large degree of autonomy. Two decades later, it could be observed that whereas some of these regional governments were highly effective and responsive to their population, other governments seemed caught in a downward spiral of corruption, lack of initiative, and poor quality of delivered services. Putnam set out to explain why some regional governments had a stronger performance record than others did. Exactly because the Italian regions had no autonomy before 1970, the Putnam study is better able to establish causal links than the Almond and Verba study because in Italy, prior differences in policy could not be responsible for the observed differences.

Based on extensive data sources and analyses, the main conclusion of Putnam's study is that the presence of a civic political culture is the main determinant of government performance. In regions with a vibrant civic culture, regional governments are more likely to respond in an effective manner to citizens' demands. Putnam identified a number of elements of this political culture. First of all, he considered the presence of *voluntary associations* to be not only a structural component of social capital. According to Putnam these associations not only dispose members to be socialized into a more socially oriented value pattern but also allow citizens to establish collective goals in a more effective manner. Following political news in the mass media is also considered an important indicator of a civic culture: If citizens read newspapers, they are likely to acquire political information, and they have the means to hold politicians accountable for policy decisions and outcomes.

Putnam's research shows that in Italian regions where these elements of a civic culture are present, regional governments perform much better than elsewhere. The study remains rather vague, however, about the precise causal mechanisms involved. On one hand, it can be expected that because of the active interaction among citizens, democratic political attitudes are being interiorized, and the population as a whole becomes more closely involved in the way society functions. On the other hand, however, it can also be expected that political elites simply experience more pressure from public opinion, forcing them to react in a more responsive manner to demands from the population.

Putnam assumes that the origins of the present-day political culture in the various regions of Italy have to be traced back several centuries. Already in the 14th century, the city-states in the northern part of Italy had established basic forms of self-rule, giving political power to some of the most privileged groups of society. The argument is that this practice had already instilled some form of democratic awareness, so citizens had the feeling that they themselves were responsible for the way their political system was being run. Southern Italy, on the other hand, during that period was still being run by autocratic monarchs, which instilled a form of subject political culture. Even six centuries later, it is argued, this division is still present in contemporary Italian society.

Empirical Research

Numerous empirical studies are available on the development and the consequences of political culture. Most often, these studies tend to focus on one specific aspect of political culture: political efficacy, political trust, political interest and knowledge, and political participation.

Political Efficacy

Political efficacy is the expression of a feeling of empowerment with regard to the political system. It is an

important attitudinal component of political culture because efficacy can be considered a prerequisite for any form of political participation to occur (Finkel, 1985). Efficacy can mainly be understood as a form of political empowerment: Citizens have the feeling that their opinion matters, that they are qualified to have an opinion on politics, and that if they make an effort to get their voice heard, they will have an impact on the decision-making process. Furthermore, political efficacy can be seen as a self-reinforcing attitude: Simply taking part in participation acts has a positive effect on the development of efficacy.

Measurement of political efficacy tends to make a distinction between internal political efficacy and external political efficacy. *Internal political efficacy* refers to an internal locus of control, that is, the feeling that the individual has sufficient (cognitive) skills and resources to arrive at a full understanding of what goes on in the political system. *External political efficacy*, however, is just as crucial. This dimension refers to the expectation that political decision makers will act in a responsive manner to the way citizens express their demands. The two dimensions of political efficacy do not necessarily go hand in hand: It is possible for a citizen to have the firm conviction that he or she is quite capable of expressing well-informed political opinions but to believe simultaneously that authoritarian rulers will not pay any attention to what citizens say. However, survey research often shows that in practice, internal and external political efficacy are closely related.

Political Trust

Political trust (sometimes also called *institutional trust*) refers to the feeling that citizens have of the trustworthiness of political institutions. As such, it can be seen as a form of diffuse support a system receives from citizens. It is important in this regard to make a distinction between different levels of the object of political trust. First, political trust can refer to the conduct of specific politicians, most notably those holding power. Unsurprisingly, this kind of specific trust is most volatile, and it is related to the performance of the politicians in power, of the political system in general, and of the economy of the country. Fluctuations in the level of trust toward specific politicians, however, are usually not seen as problematic, as incumbent politicians can be replaced.

A second level refers to trust in the institutions themselves, such as parliament, government, or the courts. Here the assumption is that trust in these institutions is vital for the stability of the system. Even if citizens do not agree with the policies that are being conducted by the governing party or parties, they may yet express trust in the institutions of government. Survey research indicates that trust in institutions is much more stable than trust in specific politicians. Third, and much more abstract, is trust in general principles governing political life, such as trust in democracy. In most Western societies, this trust level remains very elevated and stable.

In survey research, political trust is usually measured by asking respondents to indicate which, from a list of specific political institutions, they trust. In most of the available surveys, political trust can be considered to be one-dimensional, indicating that the scores on the different institutions will be related. In other words, respondents who have trust in parliament are also likely to express trust in, for example, the police force. Some other studies, however, reveal a dual structure, with a distinction between representative institutions (e.g., Parliament, government) on one hand and nonelected, law-and-order institutions (e.g., courts and police) on the other.

The recent literature contains quite a bit of discussion of whether political trust really should be seen as a valuable—or even indispensable—aspect of a democratic political culture. It has also been argued that in democratic regimes, citizens should adopt a critical attitude toward political decision makers, and that they should not put blind trust in the political elite (Norris, 1999). Some authors have even argued that it should be assumed that citizens will be distrustful toward the political elite, because there is no reason to expect that politicians will be motivated first of all by the interests of the people.

Research shows, however, that the presence of political trust has considerable consequences for the effectiveness of the political system. Scholars have noted a strong distinction between countries with high levels of political trust (most notably the Scandinavian countries) and the former authoritarian regimes in Eastern Europe (Newton, 2007). High levels of political trust are associated with greater willingness of citizens to obey the law, with a higher level of legitimacy for state intervention, and with a reduced need for government control on, for example, tax legislation (Tyler, 2006). Here too, however, the direction of causality is uncertain. One can indeed observe a close relationship between levels of corruption in a country and the level of political trust among its citizenry. Studies on Central and Eastern Europe, however, have established the reverse causal relationship: Endemic corruption in a country has a negative impact on the development of political and institutional trust in the population (Mishler & Rose, 2001).

The United States offers a specific case in the study of political trust. Although in the 1950s, the United States clearly belonged to a group of countries with very high levels of political trust, the level of trust has continuously declined since the late 1960s. In the early 1970s, the Watergate affair led to a spectacular drop in trust levels, and the level of trust has never been restored to historical levels (Hetherington, 2005). It is difficult to determine the exact cause of this structural decline, but a number of arguments have been put forward. First, increasing ideological polarization between the political parties has rendered it more difficult to respect the opinions and the political functioning of opponents. Second, the way in which the mass media cover political affairs has become increasingly aggressive and cynical. During election campaigns, for

example, the media place much more emphasis on strategic considerations and the “horse race” coverage of campaign dynamics than on the ideological and political differences between the candidates. Various scandals and affairs also receive considerable attention in the media, leading to the perception that corruption might be widespread in Washington, DC. The decline in levels of political trust also has political consequences. Hetherington (2005) argues that U.S. citizens are reluctant to accept government intervention in various fields (e.g., health care or education) exactly because they do not have the feeling that state institutions are able to deliver those services in a reliable or cost-effective manner.

Political Interest and Knowledge

The civic culture also requires that citizens be actively interested in political developments and that they expend time and effort to keep abreast of political developments. Cognitive involvement is thus an essential component of the civic culture. There is more discussion in the literature, however, on how exactly this form of involvement should be conceptualized and hence operationalized. Early research from the 1950s argued that citizens should acquire knowledge about the functioning of the political system, and therefore the state of political knowledge among the population could be used as an indicator of the presence of a sound political culture. Routinely, however, political knowledge levels in the population were shown to be very low, with most citizens apparently unaware of even the basic facts with regard to the functioning of the political system, and this result was of course a reason for concern among numerous political science scholars. During the 1970s and 1980s, however, this view was increasingly abandoned, as it was shown that political knowledge levels were mainly determined by education levels. Some deduced from this finding that those with lower education levels somehow could be blamed for not contributing sufficiently to the level of political knowledge (and hence of civic culture).

However, a number of studies by Michael Delli Carpini and Scott Keeter (1997) put political knowledge on the research agenda of political scientists. Carpini and Keeter's study demonstrated that cognitive mobilization has strong consequences for various forms of political engagement. Those high on political knowledge are not only more tolerant, even toward citizens with other political preferences, but are also better able to identify their own political preferences and the way they match with the positions taken by political parties and politicians. To put it more simply, political knowledge allows citizens to participate in politics in an effective manner. Contemporary concern was further fueled by the finding that mass media play an increasingly divisive role with regard to political knowledge. Although entertainment media have a negative correlation to political knowledge, highbrow news media add to the development of political knowledge (Prior, 2005).

This impact, of course, means that the knowledge gap between the haves (i.e., those with a high education level and a preference for news media) and the have-nots (those with low education levels and heavy use of entertainment media) will only continue to expand.

This trend does not mean, however, that political knowledge has become a generally accepted indicator of the presence of a civic culture. First of all, although it is possible to measure political knowledge in a more or less valid manner within one country, comparative tests of political knowledge (to be applied in a number of countries simultaneously) are almost totally absent from research. Even within a country, various questions have been raised about the validity of political knowledge tests. It has been shown, for example, that male respondents perform better on questions about leading male politicians, whereas female respondents respond better on questions about leading female politicians, serving as role models for female respondents, who apparently are able to identify more easily with female politicians (Mondak & Anderson, 2004).

A more indirect manner of testing cognitive involvement in the political process is to question the level of political interest of respondents. In this kind of research, respondents are simply asked whether they are interested in politics. How exactly they practice or express this political interest is not asked. This kind of survey question seems to work very well: Political interest is not only a stable attitude but also has a strong predictive effect on reading newspapers, on following political news on television, and on the level of political knowledge.

Political Participation

Citizens are also expected to participate in political life in order to convey information about their demands and preferences to political decision makers. Although various definitions of the concept of political participation can be found in the literature, they all have two elements in common. First, political participation is a form of action to express a demand. Whether citizens go out to vote, take part in a demonstration, or write an e-mail to a member of parliament, it is clear that they have a specific preference or demand and are asking the political system to respond in some way or another. Second, they are able to put some pressure on the decision makers to pay attention to their demand. This is most clear in elections, as it can be assumed that politicians who do not pay sufficient attention to the preferences of the population will not be returned to office. Demonstrations can also be used to disrupt the normal functioning of society or to endanger the legitimacy of government. Although the element of pressure might be weaker in an act such as writing to a member of Congress, the implicit threat is still that not responding to the demand will lead to fewer votes at the next election (Rosenstone & Hansen, 1993).

Political participation is determined, to a large extent, by the presence of resources (Verba, Scholzman, & Brady,

1995). The *civic voluntarism model* argues that citizens will participate only if they have time, money, cognitive skills, and other resources available. Some persons are also more likely to be targeted by mobilization efforts than are others. The model also includes *civic skills*, such as the ability to address a meeting, to voice a concern in a coherent manner, or to discuss politics with others. This model implies that strong differences in level of political participation occur. Specifically, citizens with higher levels of education will participate more often, more intensively, and more effectively than those with lower levels of education. Verba in particular claims that this form of inequality can lead to strong distortions in the political process: Those with abundant political resources will enjoy more opportunities to get their voices heard in the political process than those who have less access to these resources (Verba, Nie, & Kim, 1978).

In practice, citizens have access to a whole array of possible participation acts, ranging from voting to party activism to illegal protest. In the research on political participation, the repertoire of acts that are considered legitimate participation acts has gradually widened. In the 1950s, the emphasis was still strongly on voting in elections, taking part in electoral campaigns, or activities within political parties. In their hallmark study on *Political Action*, Barnes, Kaase, and Allerbeck (1979), however, argued that one should distinguish two totally different kinds of political participation. On one hand, conventional political participation refers to activities taking place within the context of mainstream political institutions such as political parties. But in addition, citizens have access to various other forms of participation that are much more elite-challenging. Examples would be signing petitions and taking part in demonstrations or even illegal protests. Barnes et al. show clearly that taking part in these unconventional forms of political participation is not an indication of total alienation from the political system. On the contrary, citizens who take part in one form of participation are also more likely to take part in other forms. Party members might, if the need arises, take part in demonstrations, for instance. Barnes et al. conclude that both conventional and unconventional participation acts have their place in a democratic political system and that citizens apparently select the participation act that they consider to be most effective in specific circumstances.

Although the 1979 Barnes et al. volume has been immensely successful, in the current literature it is also considered outdated to some extent. First of all, the distinction between conventional and unconventional participation has become blurred. Elite political actors increasingly rely on unconventional means to get their voices heard. Second, various new forms of political activism are gradually becoming more important in the early years of the 21st century. Internet activism is one obvious example. The Internet is increasingly being used as a medium to express political opinions (Krueger, 2002). Another addition to the political action repertoire is the rise of political

consumerism, or consumers' use of all kinds of boycott campaigns to get their message across, not just to national governments, but to international organizations or international corporations (Stolle, Hooghe, & Micheletti, 2005). It is unclear whether these "new" forms of political participation should be labeled conventional or unconventional participation.

Developments in Contemporary Political Culture

As mentioned earlier, the civic culture approach to the study of political culture met with increasing criticism from the 1970s on. Various authors questioned Almond and Verba's focus on loyalty and deference to authority in their notion of a civic culture. Increasingly, authors argued that citizens should become more active, and should enjoy a larger freedom to express their political opinions. Ronald Inglehart is one of the most prominent figures in this line of research. From the early 1970s, he argued that a gradual *cultural shift* is taking place in Western societies. Increasingly, citizens develop a *postmaterialist*, or *post modern*, value pattern. As material needs can now be taken for granted for a vast majority of the population, citizens increasingly develop postmaterial value preferences, paying more attention to quality-of-life issues, equality, and protection of the environment. One of the defining characteristics of this value pattern is a more positive appreciation of individuality and the need for self-expression. Citizens with this value pattern are less likely to respect authorities, and they demand more strongly that their individual opinions and patterns be paid attention to in the decision-making process. Whereas Almond and Verba feared that this increasing use of political participation might lead to an overload of the political system, Inglehart (1997) assumed that political systems simply have to find ways to deal with the increasing volume of demands from their citizens. The presence of critical citizens, therefore, does not pose a problem for the stability of democratic society. Inglehart produces empirical evidence that in the most stable democratic systems of the world (e.g., the Scandinavian countries or Canada), postmodern values are the most widespread across the population.

Other authors, too, have put forward the claim that cultural changes in Western societies have led to a different political culture from the one that prevailed when Almond and Verba conducted their study. Group identities have become less salient as processes of individualization have led to a more individualized outlook toward the political system. Class and religious cleavages, therefore, are less able to predict political preferences and political behavior. The feeling of loyalty toward political leaders has been eroded, and citizens now tend to develop a more critical attitude toward the political system. Furthermore, there is a strong demand for more effective ways of participating in political decision making. All these developments

impose a strain on the functioning of political systems, which are forced to invent new ways to meet these social demands. It would be wrong, however, to consider these developments as a crisis for democracy. Survey research shows that citizens support democratic government more strongly than ever before. At the same time, however, they are increasingly critical for the political institutions that have to embody this form of government (Dalton, 2004).

Future Directions

The study of political culture remains as vibrant as ever in political science, although it has to be mentioned that little progress is being made in the perennial debate between cultural and institutional perspectives on the development of political culture. For empirical scholars, it is indeed very difficult to come up with a convincing and feasible research design that would settle this debate conclusively. Nevertheless, the distinction between cultural and institutional explanations is crucial for two hotly debated topics: Can a democratic political culture be established, and can the institutions of democratic political systems adapt to the rise of a new generation of "critical citizens"?

One of the main problems in the study of political culture remains to define how exactly a democratic political culture can be installed. Especially in the former authoritarian regimes of Central and Eastern Europe, the lack of support for democratic principles of government remains highly problematic. Even 20 years after the fall of the authoritarian regimes, it can be observed that levels of political trust, political interest, and political efficacy remain remarkably low. Although in some countries, an upward trend for these indicators can be observed, in other countries the levels remain very low. An institutional approach would argue that trust levels will rise only if the political institutions themselves will start to function in a more effective manner, by reducing the current high levels of corruption. In this regard, too, however, there is little reason for strong optimism: It cannot be expected that a fight against corruption will lead to a quick or immediate rise in political trust levels. Cultural approaches argue that the relation should be seen the other way around: Only if the political culture of the population changes in a profound manner will it become less likely that civil servants or politicians will resort to corruption. The basic question in this regard is how stable political cultures are. In his famous book *The Clash of Civilizations*, Samuel Huntington (1996) argues that political cultures are inherently stable and that they do not change much over time. The interaction between cultures, therefore, would easily lead to enduring conflicts. Empirical research, however, shows quite convincingly that cultures can and will change over time. One example may be instructive: In the early 1950s, there was considerable concern over whether the political culture in Germany was conducive to stable democracy. Little more than a decade later, however, survey research showed that the German population in a very

stable manner supported democratic orientations and values, and Germany is now considered to be one of the most stable democracies in the world. Why exactly this transition occurred, and how this experience might be transferred to other social contexts, however, remains a question that needs to be explored further.

Another important debate concerns the impact of the rise of a new generation of critical citizens. That citizens in Western democracies have become more critical has been well established. Given the fact that this attitude is concentrated among younger and more highly educated cohorts, it can also be reasonably expected that this trend will continue to grow in importance in the decades ahead. What we do not know yet, however, is what the impact of this trend will be for the functioning of political systems. Is this just another challenge that democracy can easily overcome (Dalton, 2004)? In the framework of David Easton (1965), it is assumed that political systems are dependent for their stability and legitimacy on the diffuse support they receive from the population. We do not know what will happen if this form of loyalty becomes weaker. To put it differently, is it possible that citizens become too critical? Electoral research shows that a lack of political trust can lead to a vote for extremist or populist parties, and this too might form a threat for the stability of democratic political systems.

The rise of the generation of “critical citizens” poses a much more fundamental question for political science, and more broadly for the social sciences. On one hand, we know that almost all indicators of a democratic political culture are closely related to the education level of respondents. Those who have high educational levels are usually more trusting, and they also have a stronger feeling of political efficacy. Simultaneously, we know that the average education level of citizens in Western countries has risen strongly since the mid-20th century. So the assumption could be, for example, that feelings of efficacy would have risen in the same manner during this period. However, it is clear that this has not happened. Although access to higher education has been generalized in most Western democracies, the expected results of this evolution sometimes fail to materialize.

Conclusion

Within democratic political systems, citizens’ attitudes toward the political system clearly matter. Traditionally, it was expected that citizens trust the political institutions but that they also feel sufficiently efficacious to play a part in political life and to get their voice heard. While there is a general consensus on the importance of a democratic political culture, there is an ongoing discussion about the direction of causality: Does a democratic political culture lead to democratic stability, or do effective institutions lead to the development of a democratic political culture? This problem largely remains unsolved.

Empirical research, however, shows that political trust levels tend to decline, especially in the United States. How exactly political institutions retain their legitimacy and effectiveness in these circumstances remains to be investigated. A compelling issue, furthermore, is how a democratic political culture can be built in newly democratic systems.

References and Further Readings

- Almond, G. A. (1956). Comparative political systems. *Journal of Politics*, 18(3), 391–409.
- Almond, G., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.
- Barber, B. (1984). *Strong democracy: Participatory politics for a new age*. Berkeley: University of California Press.
- Barnes, S., Kaase, M., & Allerbeck, K. R. (1979). *Political action: Mass participation in five Western democracies*. Beverly Hills, CA: Sage.
- Cain, B., Dalton, R. J., & Scarrow, S. (Eds.). (2003). *Democracy transformed? Expanding political opportunities in advanced industrial democracies*. Oxford, UK: Oxford University Press.
- Dalton, R. J. (2004). *Democratic challenges, democratic choices: The erosion of political support in advanced industrial democracies*. Oxford, UK: Oxford University Press.
- Dalton, R. J. (2007). *The good citizen: How a younger generation is reshaping American politics*. Washington, DC: CQ Press.
- Delli Carpini, M., & Keeter, S. (1997). *What Americans know about politics and why it matters*. New Haven, CT: Yale University Press.
- Easton, D. (1965). *A systems analysis of political life*. New York: Wiley.
- Easton, D., & Dennis, J. (1960). *Children in the political system: Origins of political legitimacy*. New York: McGraw Hill.
- Finkel, S. (1985). Reciprocal effects of participation and political efficacy: A panel analysis. *American Journal of Political Science*, 29(4), 891–913.
- Flanagan, S. C. (1982). Measuring value change in advanced industrial societies: A rejoinder to Inglehart. *Comparative Political Studies*, 15(1), 99–128.
- Geddes, B. (1999). What do we know about democratization after twenty years? *Annual Review of Political Science*, 2, 115–144.
- Harrison, L. E., & Huntington, S. P. (Eds.). (2000). *Culture matters: How values shape human progress*. New York: Basic Books.
- Hetherington, M. (2005). *Why trust matters: Declining political trust and the demise of American liberalism*. Princeton, NJ: Princeton University Press.
- Huntington, S. (1996). *The clash of civilizations and the remaking of world order*. New York: Simon & Schuster.
- Inglehart, R. (1988). The renaissance of political culture. *American Political Science Review*, 82(4), 1203–1230.
- Inglehart, R. (1997). *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*. Princeton, NJ: Princeton University Press.
- Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change and democracy: The human development sequence*. New York: Cambridge University Press.

- Jennings, M. K. (1996). Political knowledge over time and across generations. *Public Opinion Quarterly*, 60(2), 228–252.
- Jennings, M. K., & van Deth, J. W. (1990). *Continuities in political action*. New York: Walter de Gruyter.
- Karp, J., & Banducci, S. (2008). Political efficacy and participation in twenty seven democracies: How electoral systems shape political behaviour. *British Journal of Political Science*, 38(2), 311–334.
- Klingemann, H. D., & Fuchs, D. (Eds.). (1995). *Citizens and the state*. Oxford, UK: Oxford University Press.
- Krueger, B. S. (2002). Assessing the potential of Internet political participation in the United States. *American Politics Research*, 30(5), 476–498.
- Mishler, W., & Rose, R. (2001). What are the origins of political trust? Testing institutional and cultural theories in post Communist societies. *Comparative Political Studies*, 34(1), 30–62.
- Mondak, J., & Anderson, M. (2004). The knowledge gap: A reexamination of gender based differences in political knowledge. *Journal of Politics*, 66(2), 492–512.
- Muller, E., & Seligson, M. (1994). Civic culture and democracy: The question of causal relationships. *American Political Science Review*, 88(3), 635–652.
- Mutz, D., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, 99(1), 1–15.
- Nevitte, N. (1996). *The decline of deference*. Peterborough, ON, Canada: Broadview Press.
- Newton, K. (2007). Social and political trust. In R. Dalton & H. D. Klingemann (Eds.), *Oxford handbook of political behavior* (pp. 342–362). Oxford, UK: Oxford University Press.
- Norris, P. (Ed.). (1999). *Critical citizens*. Oxford, UK: Oxford University Press.
- Parry, G., Moyser, G., & Day, N. (1992). *Political participation and democracy in Britain*. Cambridge, UK: Cambridge University Press.
- Pharr, S., & Putnam, R. (Eds.). (2000). *Disaffected democracies: What's troubling the trilateral countries?* Princeton, NJ: Princeton University Press.
- Prior, M. (2005). News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. *American Journal of Political Science*, 49(3), 577–592.
- Putnam, R. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Putnam, R. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- Rosanvallon, P. (2008). *Counter democracy: Politics in an age of distrust*. Cambridge, UK: Cambridge University Press.
- Rosenstone, S., & Hansen, J. (1993). *Mobilization, participation, and democracy in America*. New York: Macmillan.
- Schudson, M. (1997). *The good citizen: A history of American civic life*. New York: Martin Kessler Books.
- Smith, G. (2009). *Democratic innovations: Designing institutions for citizen participation*. Cambridge, UK: Cambridge University Press.
- Stolle, D., Hooghe, M., & Micheletti, M. (2005). Politics in the supermarket: Political consumerism as a form of political participation. *International Political Science Review*, 26(3), 245–269.
- Tocqueville, A. de. (2007). *Democracy in America*. New York: Norton. (Original work published 1835–1840)
- Tyler, T. (2006). *Why people obey the law* (2nd ed.). Princeton, NJ: Princeton University Press.
- Verba, S. (2003). Would the dream of political equality turn out to be a nightmare? *Perspectives on Politics*, 1(4), 663–679.
- Verba, S., Nie, N., & Kim, J. (1978). *Participation and political equality: A seven nation comparison*. Cambridge: Cambridge University Press.
- Verba, S., Schlozman, K. L., & Brady, H. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

25

RELIGION AND COMPARATIVE POLITICS

BRYAN A. S. RASMUSSEN

Collin College

The past two decades have seen a resurgence in the study of how religion affects politics in the United States and around the world. For generations, social scientists believed religion to be declining in influence to the point that it might eventually be marginalized. However, political scientists continue to observe, among other things, the importance of Christianity in the United States and the increasing influence of extremist Islam leading to events such as September 11, 2001. As political scientists have asked questions about these developments, the body of literature on the subject has grown to the point that the American Political Science Association recently initiated a journal titled *Politics and Religion* in order to give proper attention to this important area of research.

This chapter will review contributions scholars have made to the understanding of religion and politics from a comparative perspective. Although the United States is one case that is worth examining because it is unique, the chapter will touch on it only briefly. There is another chapter in this volume that deals specifically with the relationship of religion and politics in the United States. Analyzing religion and politics from a comparative perspective makes the analysis a great deal more complex. There are a wide variety of government institutions and cultures around the world, as well as a variety of religions represented. In addition, there is considerable distinctiveness within many of

those faiths as to how the sacred texts are interpreted or how “faith” is to be practiced.

There is also considerable variation in the methods comparative political scientists use to do their research. The days of comparativists doing exclusively qualitative area studies are gone. While some of that valuable research remains, comparativists now have the data and computing capacity to strive for more quantitatively supported grand theories. The problem that has been seen in the past, as this chapter shows, is that the oversimplification that is required to make a grand theory can many times render the theory weak or simply wrong. In the end, the prevailing wisdom among comparativists calls for building theory from the bottom up (Geddes, 2003). This means researchers start with a country or a region, establish a relationship there, and then try to generalize as they can from there. This means that the variations in institutions, cultures, and religions previously mentioned cannot be ignored, making religion and comparative politics a very complex subject.

As such, this chapter does not cover the whole topic. Instead, it starts by examining the fundamental role of religion in society and how that role has evolved over the years. It then reviews some of the primary approaches scholars have taken to understanding the relationship between religion and politics around the world. Finally, it looks specifically at some of the most important areas of study in the field in recent decades.

The Evolution of Religion Theory

To understand the current state of the literature in the field of religion and comparative politics, it is helpful to begin with the influence religion can have on individuals. It goes without saying that religion has always been important to people and has the capacity to influence how they live their lives. For many, religion is a cornerstone of who they are and plays a big part in many of the decisions they make. As will be discussed, there is evidence that religion has been losing influence over the past 400 years and that this pattern continues today, although the extent to which this fact is true is up for debate. In any case, religion has always had, and continues to have, a very personal connection to many people and can therefore be used in many ways as a powerful motivator.

From where does this deep personal connection come? Social scientists have been debating many possible answers to this question for centuries. In *Eight Theories of Religion*, Daniel Pals (2006) identifies primary modern theories and traces them back to their origins in an effort to understand how this connection between people and religion is formed and why it is so important. Before a few of the social scientists that Pals credited with defining these theories are highlighted, two points need to be made. First, it is interesting to note that most of these early social scientists took general theories that they had been working on and applied them to religion. They had a lens through which they saw the world, and they then looked at religion through that lens, giving us a different perspective.

Second, most of these social scientists were not religious people. A certain religion (i.e., Christianity) was not the lens through which they saw the world. It was scientific theory that provided their perspective. One of the assumptions sometimes made in discussions about the nature of religion is that devoutly religious people cannot effectively contribute to an intellectual discussion about religion. For a religious person, religion is what it is. There is an absolutism to religion that would make it difficult, if not impossible, for a religious person to look at religion from an intellectual viewpoint. However, the point can be made in the other direction as well. Can an irreligious social scientist properly understand the impact a religion has on a person or a culture without properly understanding the supernatural aspects of the religion itself? This debate about the role of social scientists in analyzing religion is one that continues in political science journals today (see Mitchell, 2007; Wilcox, Jelen, & Wald, 2008).

Pals begins by discussing two anthropologists, E. B. Taylor (1871) and J. G. Frazer (1911–1915), who endeavored in different ways to discover the role of religion in “primitive” societies. They both felt that this investigation would provide insight into how religion developed into such an integral part of the different cultures around the world. Two main themes come from Pals’s discussion of

Taylor and Frazer. First, humans have a basic desire to believe in something greater than themselves. This *animism* can be found contributing to religious traditions as well as other mythological traditions.

The other theme was the desire to explain the seemingly unexplainable. For these societies, much of the world could not be explained. To say these things were magic was one possible explanation, but to form a religion around these events and objects provided an opportunity to rationally explain the unexplainable. This would in turn give order and structure to the culture, as those who understood the “gods” would be granted positions of authority. *Divine right* would become the primary source of legitimacy for governments around the world for centuries, and it still is a source of legitimacy for some governments today. For Christianity, becoming the official religion of the Roman Empire about the year 400 CE gave it an authority that would go nearly unquestioned in many ways for more than 1,000 years.

Although these anthropological studies provide certain insights, they do not directly explain the influence of religion in most of today’s world. Although there are still pockets of “primitive” societies around the world, most of the world has developed far beyond that level. The biggest historical step in this development was arguably the printing press. This enabled people to have access to information themselves. In terms of Christianity, a big reason for the Church’s authority was that no one could read the Bible, including many priests. Many people today, Christians and non-Christians alike, accept the fact that many parts of the Bible can be interpreted many different ways. Prior to the printing press and the Gutenberg Bible, there was no opportunity for individual interpretation.

Among the first wave of challengers to the authority of the Christian Church was Niccolò Machiavelli. Although he was subtle and respectful in his approach (as he had to be at the time in order for anyone to listen to him), it was clear he was challenging the Church as an absolute authority. This challenge would be followed soon by Martin Luther and the Protestant Reformation. Luther, like many other visionaries, was trying only to improve the existing structure by challenging the leaders to address the corruption that had developed. Instead, he ended up inspiring a revolution that changed the face of Christianity around the world.

The power of the exchange of ideas would soon move beyond religion and inspire the Enlightenment. This movement of liberal philosophers spanned the 16th and 17th centuries and represents an important turning point in our discussion. Although John Locke and many of these thinkers still wrote in a religious context, so as not to alienate their potential readers, they wrote about a more secular world. Locke and Montesquieu envisioned a world where commerce was central to people’s lives and secular education was to be of vital importance. They felt people had the ability to reason, and therefore they should

be given freedom to run their own governments (via democracy) and invest themselves in science or commerce as they saw fit. These philosophies, first incorporated by the founders of the United States and since throughout most of the world, have led to industrial revolutions and centuries of unprecedented advancement in science and technology. However, what does all this mean about the role of religion?

Religion had begun to lose its absolute authority in many societies. In addition, science and technology were beginning to provide answers to questions that previously could be answered only by religion. Friedrich Nietzsche, a German philosopher critical of the Enlightenment, went as far as saying that he envisioned a world where God would be dead. Science may not answer all the questions, but if commerce is there to keep us preoccupied, then we may lose our drive to delve into deep issues in which religion could still provide guidance. The bottom line of this view is that religion could well lose its relevance as a significant political force.

This theory gained influence among social scientists of the late 19th and early 20th centuries. Pals reviews the works of three men who reduce religion to an expendable role in their view of the world. Sigmund Freud (1913/1953) saw religion as a form of psychological neurosis that could be eliminated with proper treatment. For Émile Durkheim (1915), religion provides a sense of social order and belonging, things that could potentially be achieved by other means. Karl Marx (1867/1990) saw religion as an illusion that was propagated to maintain the class structure. Each author fit religion into his lens, and each saw it as expendable as a political force.

For the remainder of his book, Pals focuses on social scientists, including Max Weber, who seemed to have a better appreciation for the complexity of religion. Weber's approach to religion and politics is more direct and complex than the social scientists previously mentioned. His signature work, *The Protestant Ethic and the Spirit of Capitalism* (1930), places religion at the center of who we are as a people and what we do. However, at the same time, he says that once these behavior patterns become ingrained in us, they are no longer tied to the religion, making the religion less significant. Therefore, despite his belief that religion is quite important to the formation of cultural values, Weber, among other things, contributed to the growing notion that religion may one day be marginalized.

The Secularization Theory

The dominant paradigm in the social sciences through most of the 20th century centered on the idea that religion was headed toward extinction as a political force. Eventually, this idea was formalized into the *secularization theory*. Peter L. Berger (1969), one of the leading proponents of the theory, said that reason, scientific development, and bureaucratic specialization were among the factors that would

eventually destroy religion as a political influence. As this grand theory was the dominant paradigm for so long, not much was written about religion and comparative politics for decades except for pieces supporting the secularization theory or religion's being an aspect of the qualitative analysis of an individual country's culture. Communism was spreading through the world without the use of religion (overcoming religious obstacles, Marx might say). At the same time, Western Europe was becoming less religious by the decade. The pattern of secularization worldwide, at least in more developed nations, seemed to be clear.

This all began to change in the 1970s. Political scientists began to ask questions about two main trends that seemed to be challenging the secularization theory. First, although religiosity was on the decline in Western Europe and much of the "first" world, this pattern was not as strong in the United States. The second trend was the *liberation theology* that spread through Latin America at that time. In the late 1960s, the Catholic Church held a council called *Vatican II*, resulting in documents that, among other things, encouraged Catholics to be more active in expressing their faith. Latin American Catholics, much to the dismay of the Vatican, interpreted this to mean that they should band together as Catholics and rise up against the oppression of the authoritarian governments that were common in Latin America at the time. This use of religion to motivate a bottom-up political movement was unique and provided evidence that religion was still very much alive as a political force. (For more on liberation theology, see Gill, 1998; Levine, 1986a.)

An article by Daniel Levine (1986b) shows the beginnings of a trend to counter the secularization theory. It is basically a review of books written in the early 1980s that shed new light on changes in religions and their roles in societies. Later, in the introductory chapter to his collection of works titled *The Desecularization of the World*, Berger (1999) recants his previous support for the secularization theory. Berger goes on to say that modernization can have the opposite effect of actually promoting religion. This countersecularization movement can be driven by the people (conservatives) or by elites searching for expanded legitimacy. The bottom line is that the secularization theory, as it stood into the 1980s, is now dead, and the question becomes, Where do we go from here?

Scope of Theory

As has been shown, the prospect of developing a grand theory, in this case one theory that explains the relationship between religion and politics all over the world, is problematic. Freud's and Marx's attempts to incorporate religion into their grand theories left us with more questions than answers. With the demise of the secularization theory and the spread of the bottom-up theory, it might not make sense even to attempt another grand theory. Yet the temptation is always there for the social scientist. There have been two

attempts at a grand theory of religion and comparative politics since 1990 that are worth mentioning here.

In *The Clash of Civilizations* (1996), the late Samuel Huntington sees religion as a major aspect of global realignment after the cold war. The end of the Soviet Union saw the end of 45 years of a global bipolar balance of power. The world had taken sides, and the battle between capitalism and communism dominated the world stage. The question Huntington is trying to answer is, Now what?

In Huntington's vision, religious and other ethnic factors divide the world into two main categories: The West and the rest, or Christian democratic nations (and a few others for strategic reasons) versus the rest of the world (of which the primary unifying force is Islam). Huntington discusses the decline of the, until now, dominant Western culture leading to opportunities for regional conflict to escalate along these fault lines. In his conclusion, he paints a scary scenario about the potential for World War being driven from these divisions, although this scenario is not presented as inevitable, just as an example of how his vision of the world could go wrong.

A counterparadigm for the role of religion in the world comes from *Sacred and Secular*, by Pippa Norris and Ronald Inglehart (2004). They argue that the secularization theory is not a relic; it just needs some modification to properly explain today's world. The theory still applies to the modern industrialized world, taking into account historical differences, but not to the rest of the nations of the world, which have yet to modernize and replace religion as the focal point of their society.

This work is a largely quantitative study using data from the World Values Survey (www.worldvaluessurvey.org/). Basically, the authors apply Inglehart's postmaterialist theory to the role of religion in the world. They say that developed nations get more secular as they get more secure. As a result, postindustrial countries have less religious alignment and lower church attendance. In addition, birth rates are lower in these countries, and therefore Norris and Inglehart see religion actually growing in the world because population rates are increasing faster in the less developed world, which intensifies the line of conflict between the secure world and the third world.

Norris and Inglehart paint a picture much different from Huntington's, and they present data that challenge Huntington's vision. However, there are anomalies in their theory, the primary one being the United States. They explain that the minimal U.S. social welfare system fosters enough economic insecurity in our cities and rural areas to cause religiosity similar to that in areas of the third world. This explanation seems weak because one could easily argue that the homeless person living in a shelter in Los Angeles has a more comfortable life than half of the population of a country like Bangladesh.

While both of these attempts at a grand theory are important, they are arguably oversimplified and incomplete. As such, most of the literature on religion and comparative politics over the past 30 years has been smaller

in scope. Researchers analyze one country or region in depth to try to show a pattern or relationship between variables. They might also compare two countries or regions and try to explain a similarity or a difference between them. Once these connections have been made, it may be possible to generalize the theory from there, but even if not, it is still a valuable contribution to our understanding of the relationship in a particular context. This research is also helpful for comparativists who had previously done qualitative work in a specific area of the world. They can still focus on their region of expertise as they add quantitative analysis.

An excellent example of this regional approach is a compilation of chapters edited by Ted G. Jelen and Clyde Wilcox (2002) titled *Religion and Politics in Comparative Perspective*. Contributors to this book cover issues in religion and politics in countries from Japan to India and Spain, just to name a few. Researchers have written articles of note about Australia (Bean, 1999), China (Weiming, 1999), Israel (Wald, 2002), and many countries around the world. However, there are a few regions that have gotten the bulk of the attention because they provide clear examples for some of the primary theories being currently tested. These countries and regions include the United States, Europe, Latin America, and the Islamic Middle East.

Primary Regions and Theories of Current Interest

Europe

Europe is a good place to start, as it has always been the model for the old secularization theory. Even today, evidence of religious dealignment and very low church attendance abound throughout most European nations, making Europe a unique study.

In the late 1960s, Berger developed the *sacred canopy model* out of the secularization theory and used it to describe these European countries. In those days, many of the countries had a formal church-state institution. Take Catholic Italy, for example. The canopy model says that because of this formal institutional link, a linkage also develops in the culture. It is as if being Italian becomes intertwined with being Catholic to the point it is difficult to distinguish whether a cultural value is distinctly Italian or Catholic. This can be evidenced by non-Catholic Italians professing Catholic values.

The church's going unchallenged has a number of ramifications. First, church attendance lowers (if you are Catholic from birth, you do not have to go to church every week to prove it). Second, a professionalization of the clergy occurs as its members find themselves less connected to their congregants. Finally, government policy and religion grow more connected, which can be positive and negative. Certainly the government stands to gain by

association with this added source of legitimacy. Does religion gain from the alliance? Some say it does by having a more mainstream role in the society. However, others say religion loses by an alliance with government because generally democratic governments strive for compromise and consensus whereas religion is about right and wrong, leaving church and state incompatible. In any case, many European countries to this day provide strong evidence for the sacred canopy model.

This picture has two major complications, however. First is the assumption that low church attendance means a lack of religiosity. Church attendance has always been the signature variable for measuring religiosity because it is data that is easily accessible. One may ask, How are the data acquired? Are they self-reported by people who want to make themselves look good? Or do they come from the church, which also has image motives? While it is not completely reliable, church attendance has always been the best indicator we have. However, with the continuing development of worldwide surveys such as the World Values Survey, this limitation is changing. We can get a more complete picture of a concept such as religiosity with questions about belief in God, belief in hell, the authoritative nature of scripture, and so forth. These new data are very important because comparing church attendance levels in Europe with those in the United States, for example, is just not appropriate without additional evidence to support an inference, whatever it might be.

The other development that is complicating the canopy picture in many European nations (there are a few, such as Ireland, where church attendance and other measures of religiosity are very high) is a trend away from church-state institutions in Christian nations. Many nations that were church-states, such as Italy and Poland, have broken that formal link, and others, such as Norway, are considering such a move. What does it mean to no longer have a formal church-state but to still have a religion with a strong majority status? Is the resulting canopy the same, or do other religions start to come in to compete and change the religious dynamic in the country? Many studies have examined this transformation in Europe (Jelen & Wilcox, 1998) and Latin America (Gill, 1998), where the Catholic Church has always held a strong position.

It should be noted that some religions and religious sects lend themselves more easily to church-states. Islam, discussed later, first comes to mind. Mohammed was a spiritual, political, and military leader. As such, any nation that follows his example can conceive of, maybe even strive for, a combination of church and state. Within Christianity, the institution of Catholicism lends itself easily to a church-state, as has been seen often throughout history. Still, Islam and Catholicism operate well independently of government in the United States, and Protestant church-state systems operate in Norway, England, and elsewhere. Does one religion operate better as a formal church-state than another religion, and what impact does this difference have on the people of the

nation? These are empirical questions that political scientists continue to explore.

Discussing institutions leads us to focus on political parties. One issue that has drawn much attention from political scientists is the “Christian Democratic” parties found throughout Europe. While these started out as Catholic parties, they no longer operate as such today. In most cases, they are relatively liberal parties focusing on poverty and social justice issues. (For more on this development, see Kalyvas, 1998.)

Latin America

The liberation theology movement in Latin America provides a good example of two main themes of research. The first is the debate over whether political institutions work primarily from the top down, or from the bottom up. This debate is central to political parties literature, among other things. The point is not that the institutions have to work exclusively one way or the other. As a matter of fact, the truth may be that sometimes they work one way and sometimes the other. However, it is important to distinguish whether the position of a politician on abortion affects public opinion or vice versa. Many political scientists take a top-down approach, trying to show the power of the elites, including religious elites, to manipulate public opinion. As mentioned before, the liberation theology movement reminded researchers not to ignore the bottom-up approach. In any form of legitimate democracy, the people still have power, and under certain conditions they may actually decide to use it. Under other forms of government, the people can take power, but to do so requires a deep commitment, something religion can provide.

The second key concept illustrated by liberation theology is the importance of civil society and social capital (for an explanation of these concepts, see Chapter 23, titled “Civil Society,” in this volume, or Putnam, 1994, 2000). Religion is in a unique position to generate social capital for two reasons. First, its institutions are in place, connecting people to each other and encouraging them to meet on a regular basis. Second, religions generally promote the volunteering of time and money in order to accomplish things that, in many cases, an individual will not benefit from directly. This selfless giving is not easy for many people and intensifies the connection among those who are willing to participate in it. It also commits people to the idea that there is something more important than themselves to work for. Religion has powerful potential for social action because the institutions to organize the people are there and religion has the power to motivate the people and put them in the right frame of mind for collective action. Whether that action is sparked by elite discourse or from the grass roots can vary.

The United States

Although a great deal of work has been done regarding religion and social capital in the United States (see Smidt,

2003), the concept of social capital is only a small part of what makes studying religion and politics in the United States fascinating. As a large country that uses a federal system of governance, the United States comprises state-to-state and regional differences that are interesting to study. In addition, the availability of survey data makes research easier for the political scientist. All these issues and others are covered in more detail in Chapter 98 of this volume, titled “Religion and Politics in America” (see also Wald, 2003).

This chapter focuses on one central issue that sets the United States apart from the rest of the world. As Europe has always been the best example of secularization, the United States was always the outlier. A high level of religiosity remains today, and the influence of religious groups in politics has actually increased since the 1980s in a country that tries to present itself to the rest of the world as a leader in secular concerns such as democracy and capitalism. What makes the United States different?

Many argue that the religiosity in the United States is best explained by the establishment clause in the First Amendment to the Constitution, which separates church and state. Because of the “religious free market,” churches find themselves competing for members. This competition creates a buzz that increases overall levels of religiosity. In an economic analogy, this is a supply-side approach (for the most complete work on this dynamic in the United States, see Finke & Stark, 1992). A large body of research deals with the religious marketplace in the United States.

One of the weaknesses of comparative politics in its early decades (1950s and 1960s) was U.S. ethnocentrism, a normative approach that assumed other countries would eventually want to be like the United States, and so social scientists should help them do that. Today, comparativists are much more empirical and appreciative of differences in culture. In this case, however, the U.S. example is very unique. In few places in the developed world are people as religious and the politics as affected by organized religious groups. Debates in the United States over issues such as same-sex marriage and abortion have strong religious overtones. Some political scientists have observed that since the 1980s the two political parties have become more divided along a religious versus secular fault line. If the marketplace model explains these trends, then it makes sense to see to what extent the marketplace model can be applied around the world.

Scholars have taken this model and applied it to Europe with limited success (Jelen & Wilcox, 1998). The supply-side approach can be used to investigate how much of a majority a sect has to have before a sort of religious canopy sets in. Recent research in Latin America has shown that countries in which the members of a single religion number more than 85% of the population represent the canopy model, and when the percentage of the population following a single religion falls below 85%, elements of the competition model begin to appear (Gill, 1998). It appears that a supply-side approach to religion and politics in other countries may help explain observed behavior.

This economic model suggests another dynamic as well: the rational choice model (see Chapter 5, titled “Rationality and Rational Choice,” in this volume). Religious people are not seen as rational actors because they do what God tells them, not necessarily what would maximize their utility according to secular scales. Incidentally, this logic illustrates a flaw with rational choice theory. If getting to heaven is maximizing utility, then doing what you understand God is asking of you is completely rational. Regardless, the point here is that using the supply-side economic model leads researchers to consider churches and religious elites as rational actors who are trying to maximize their utility or get more members. This possibility opens up opportunities for theories and research, not only in the United States but around the world (see Gill, 1998; Warner, 2000).

The Middle East

One of the biggest questions facing social scientists today is whether Islam is compatible with democracy. The overwhelming widespread trend toward democracy during the past 20 years has, for the most part, missed the Islamic world. The evidence clearly indicates democracies are considerably less likely to fight with each other, so promoting democracy, besides being a good thing normatively, would seem to have security value. What is it about Muslim countries that makes it difficult for democracy to take hold? To what extent can Iraq and Afghanistan answer this question if their democracies have been forced on them from outside?

This point should be left in. Scholars trying to determine answers have focused attention on many aspects of Islam and the Muslim people (see Tamadonfar, 2002). Many factors have been shown to effect democratization in general (Przeworski, 1996). In addition, most of these countries are known as *rentier states*, referring to a resource curse that may be the biggest factor hindering democratization in the Middle East. Of course, many of these factors may be working in concert. The complicated relationship of religion and politics does not operate in a vacuum. Outside factors are constantly influencing the outcomes, making the relationship that much more difficult to study.

Future Research

The possibilities for future research in this field are immeasurable. Religion is and will remain an important factor in politics all over the world. The growing possibilities for survey research in the Islamic world are exciting. Africa and the rest of the developing world offer opportunities. As survey data begin to emerge from all corners of the globe, political scientists can begin to explore some of the issues faced in other nations.

For example, most of the previous work done with regard to identity politics in the developing world has focused on ethnicity (see Horowitz, 2000). Although this

work is important, the focus on ethnicity can marginalize the perceived impact of religion. Religious differences are many times included with ethnic differences when appropriate; otherwise they are often not considered. There is a definite need for an equally strong focus on the religious aspects of these issues.

It is true that religion and ethnicity have much in common and, in many cases, interact a great deal. Both concepts are important to individuals and can elicit strong emotional responses. Both are unifying forces that can bring a group together and potentially define an enemy group. Both contribute to cultural and historical identities that, along with national identities, are used to define the world stage (Huntington, 1996; Jelen & Wilcox, 2002). In many cases, divisions among groups can be both ethnic and religious. Examples include divisions in the former Yugoslavia and many other places around the world. In some situations, it can be difficult to separate the impact of religious divisions from that of ethnic divisions.

Yet the concepts of religion and ethnicity can also play different roles in society. To begin with, they can have different foundations and historical significance. Although religious conflicts sometimes line up along ethnic lines (e.g., the Christian Croats and the Muslim Bosnians), there are also examples of religious conflict that overrides ethnicity (e.g., Iraqi Shi'a and Iraqi Sunni) or of ethnicities that combine under religious banners (as in Nigeria). Furthermore, religion may be more important than ethnicity to one person or group (e.g., Muslims) whereas ethnicity might be more important in another culture (e.g., Han Chinese). Most important, however, the groups are organized differently, and that organization can play a major role in how they can operate in society.

Religion generally has formal institutions in place that allow it to be well organized in society. An ethnic group is generally less well organized than a religion. An ethnic group does not necessarily have a supreme authority (unless the ethnic group is tied in with a religious group). There is likely leadership of some sort as a hierarchy is developed, especially in situations in which the groups are unranked. However, ethnic groups do not necessarily have a positive formal mechanism for membership or mobilization such as the church offers for religion (Horowitz, 2000).

From this institutional perspective, religion seems to be in a much stronger position to influence something like democratic stability in the developing world. The success of a newly formed democracy depends, to a large extent, on the ability of the new formal institutions to be perceived as legitimate by the people. The institutions that are already in place and respected in a country will be expected to provide a cue to the public on whether to accept the new system. Although similar to the endorsement of an ethnic leader, religious endorsement goes farther, for two reasons. First, an endorsement of democracy from the religious leadership could be interpreted as coming from a higher power. Second, the church has

formal institutions in place to mobilize the masses to accept democracy. Churches can align their message with the ideology of freedom and democracy, register people to vote, and perform many other functions to promote the new system.

In addition, the structure of the religious institutions can provide insight into culturally accepted norms of hierarchy. Tapping into a religious organization is tapping into the social and informal political institutions and links they might already have in place. These points reinforce the important institutional arguments that help us understand how support from religion could make the difference in certain new democracies. The bottom line is that the ultimate way a religion can promote the legitimacy of a democratic government is to participate in it.

Conclusion

The field of religion and comparative politics is as strong as it has ever been. A tremendous amount of work has been done in the past 20 years. However, there still is a lot of extremely important work to do. Most of the conflicts around the world have some sort of religious overtone. If people in the United States, for example, can gain a better understanding of how Islam operates in Islamic countries, that understanding will help break down some of the barriers that keep us apart.

References and Further Readings

- Bean, C. (1999). The forgotten cleavage? Religion and politics in Australia. *Canadian Journal of Political Science*, 32, 551-568.
- Berger, P. L. (1969). *The sacred canopy: Elements of a sociological theory of religion*. Garden City, NY: Anchor.
- Berger, P. L. (Ed.). (1999). *The desecularization of the world: Resurgent religion and world politics*. Washington, DC: Ethics and Public Policy Center.
- Durkheim, É. (1915). *The elementary forms of religious life*. New York: Macmillan.
- Finke, R., & Stark, R. (1992). *The churching of America 1776-1990*. New Brunswick, NJ: Rutgers University Press.
- Frazer, J. G. (1911-1915). *The golden bough: A study in magic and religion* (3rd ed., Vols. 1-12). London: Macmillan.
- Freud, S. (1953). Totem and taboo. In *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 13, pp. 1-161). London: Hogarth. (Original work published 1913)
- Geddes, B. (2003). *Paradigms and sand castles: Theory building and research design in comparative politics*. Ann Arbor: University of Michigan Press.
- Gill, A. (1998). *Rendering unto Caesar: The Catholic Church and the state in Latin America*. Chicago: University of Chicago Press.
- Gill, A. (2001). Religion and comparative politics. *Annual Review of Political Science*, 4, 117-138.
- Horowitz, D. L. (2000). *Ethnic groups in conflict*. Berkeley: University of California Press.

- Huntington, S. P. (1996). *The clash of civilizations and the remaking of world order*. New York: Touchstone.
- Ishiyama, J. (2007, September). *Do ethnic parties detract from or promote ethnic conflict?* Paper prepared for the annual meeting of the American Political Science Association, Chicago.
- Jelen, T. G., & Wilcox, C. (1998). Context and conscience: The Catholic Church as an agent of political socialization in western Europe. *Journal for the Scientific Study of Religion*, 37, 28–40.
- Jelen, T. G., & Wilcox, C. (Eds.). (2002). *Religion and politics in comparative perspective: The one, the few and the many*. Cambridge, UK: Cambridge University Press.
- Kalyvas, S. N. (1998). From pulpit to party: Party formation and the Christian democratic phenomenon. *Comparative Politics*, 30, 293–312.
- Levine, D. H. (1986a). Is religion being politicized? And other pressing questions Latin America poses. *Political Science & Politics*, 19, 825–831.
- Levine, D. H. (1986b). Religion and politics in comparative and historical perspective. *Comparative Politics*, 19, 95–122.
- Lijphart, A. (1999). *Patterns of democracy*. New Haven, CT: Yale University Press.
- Marx, K. (1990). *Capital: A critique of political economy* (Vol. 1). New York: Penguin Putnam. (Original work published 1867)
- Mitchell, J. (2007). Religion is not a preference. *Journal of Politics*, 69, 351–362.
- Norris, P., & Inglehart, R. (2004). *Sacred and secular: Religion and politics worldwide*. Cambridge, UK: Cambridge University Press.
- Pals, D. L. (2006). *Eight theories of religion*. New York: Oxford University Press.
- Przeworski, A. (1996). What makes democracies endure? *Journal of Democracy*, 7, 39–55.
- Putnam, R. (1994). *Making democracy work*. Princeton, NJ: Princeton University Press.
- Putnam, R. (2000). *Bowling alone*. New York: Simon & Schuster.
- Smidt, C. (Ed.). (2003). *Religion as social capital*. Waco, TX: Baylor University Press.
- Tamaddonfar, M. (2002). Islamism in contemporary Arab politics: Lessons in authoritarianism and democratization. In T. G. Jelen & C. Wilcox (Eds.), *Religion and politics in comparative perspective* (pp. 141–168). Cambridge, UK: Cambridge University Press.
- Taylor, E. B. (1903). *Primitive culture* (4th ed., rev., Vols. 1–2). London: John Murray.
- Tessler, M. (2002). Islam and democracy in the Middle East: The impact of religious orientations on attitudes toward democracy in four Arab countries. *Comparative Politics*, 34, 337–354.
- Verweij, J. (1997). Secularization as an economic and cultural phenomenon: A cross national analysis. *Journal for the Scientific Study of Religion*, 36, 309–325.
- Wald, K. D. (2002). The religious dimension of Israeli political life. In T. G. Jelen & K. D. Wald (Eds.), *Religion and politics in comparative perspective* (pp. 99–124). Cambridge, UK: Cambridge University Press.
- Wald, K. D. (2003). *Religion and politics in the United States*. Lanham, MD: Rowman & Littlefield.
- Warner, C. M. (2000). *Confessions of an interest group: The Catholic Church and political parties in Europe*. Princeton, NJ: Princeton University Press.
- Weber, M. (1930). *The Protestant ethic and the spirit of capitalism*. London: HarperCollins.
- Weiming, T. (1999). The quest for meaning: Religion in the People's Republic of China. In P. L. Berger (Ed.), *The desecularization of the world* (pp. 85–102). Washington, DC: Ethics and Public Policy Center.
- Wilcox, C., Jelen, T. G., & Wald, K. D. (2008). Religious preference and social science: A second look. *Journal of Politics*, 70(3), 874–879; comment, 880–883.

26

ETHNIC AND IDENTITY POLITICS

DWIGHT VICK

West Texas A&M University

JOHN T. ISHIYAMA

University of North Texas

The focus of this chapter is on the central concept of ethnic identity and its effects on politics. In particular, it focuses on the importance of ethnic identity politics from a comparative perspective. Although an enormous amount of literature on ethnic identity has been produced in anthropology and sociology, this chapter concentrates on the still voluminous literature in political science.

This chapter covers the general political science literature that addresses the following questions: (a) What is meant by the concept of ethnic identity? (b) How does an ethnic identity form and change? (c) What is the relationship between ethnicity and national identity? (d) What are the political consequences of ethnic identity (especially in terms of voting and political parties)? Finally the chapter turns to some future research questions that can be derived from this coverage of the literature.

What Is Ethnic Identity?

What is generally meant by such terms as *ethnicity*, *ethnic group*, and *ethnic identity*? The terms are derived from the Greek term *ethnos*, which has been generally translated to mean *nation* or a community of people who share a common language or culture. Although much of the early understanding of *ethnic groups*

treated these communities as natural (often conflating ethnicity with race, e.g., the German race, the English race), the notable 19th and early 20th century German sociologist Max Weber argued that ethnic groups were artificial and socially constructed. Essentially they were based on a subjective belief in a shared community. This belief is what created the group, and the motivation for creating a group derived from the desire for political power. This was very much in contrast to an earlier belief in the 19th century that held that sociocultural and behavioral differences between peoples stemmed from inherited traits and tendencies derived from common descent, or race. Later Fredrik Barth (1969) went even further, arguing that ethnicity was forever changing and that the boundaries of membership in an ethnic group are often negotiated and renegotiated, depending on the political struggle between groups.

Many social scientists have since noted the malleability of ethnic group boundaries. Thus, Joan Vincent (1974) has noted that ethnic boundaries often ebb and flow and have a rather mercurial nature. Ronald Cohen (1978) also pointed to the instrumental nature of ethnic boundaries, arguing that ethnicity “can be narrowed or broadened in boundary terms in relation to the specific needs of political mobilization.” This may be why descent is sometimes a marker of ethnicity and sometimes not—ultimately it depends on the political situation.

However, as Liah Greenfeld (1992) and many others have noted, this does not mean that ethnic identities are merely “imagined” and completely malleable and porous. Rather there are some objective characteristics that constrain identity. For instance, a Japanese American cannot suddenly declare herself to be an Ethiopian and be accepted as such in Ethiopia. Generally, others will not regard her as Ethiopian (because she does not “look it” or because she cannot speak Amharic or any other national language of Ethiopia). In other words, objective attributes constrain what can be subjectively imagined.

Thus the modern literature in comparative politics has used ethnicity as a concept that embraces these attributes. For instance, the notable scholar of ethnic politics Donald Horowitz (1985) has referred to the concept of ethnicity as a broad term that “easily embraces groups differentiated by color, language, and religion; it covers ‘tribes,’ ‘races,’ ‘nationalities,’ and ‘castes’” (p. 53). Much of the more recent literature also uses the term *ethnicity* in very much the same way—as a term that includes many other markers of identity (see, e.g., Chandra, 2004; Posner, 2005).

A very useful synthesis of objective and subjective views of ethnic identity, but one that is also more precise than previous broad conceptions, is provided by Kanchan Chandra (2006). She suggests a very useful definition of *ethnic identity* that draws on both subjective and objective views. She argues that “ethnic identities are a *subset of identity categories in which eligibility for membership is determined by attributes associated with, or believed to be associated with, descent*” (p. 398, italics added). By *ethnic identity*, she means a social category in which “eligibility for membership is determined by descent-based attributes” (p. 398). These attributes include both objective features and subjective beliefs. Thus attributes include the following:

[those] acquired genetically (e.g. skin color, gender, hair type, eye color, height, and physical features), or through cultural and historical inheritance (e.g. the names, languages, places of birth and origin of one’s parents and ancestors), or acquired in the course of one’s lifetime as markers of such an inheritance (e.g. last name, or tribal markings). Further, it includes attributes believed to be associated with descent, which mean attributes around which a credible myth of association with descent has been woven, whether or not such an association exists in fact. The definition thus includes both a subjective and an objective element. (Chandra, 2006, p. 399)

How Does Ethnic Identity Form?

Much of the literature in comparative politics has suggested that ethnic identity is one of the most powerful forces shaping political attitudes and mass political

behavior (Brubaker, 1992). Ethnicity is a type of group-based social identity, along with other group-based identities (such as class and clan), but ethnicity has had a particularly powerful effect on political behavior in the late 20th and early 21st centuries. Generally speaking, the early scholarship on ethnic identity formation was characterized by two broad debates (Eriksen, 2001), or what might be referred to as the primordial versus the situational perspectives.

The first approach (*primordialism*) generally holds that ethnic identity is innate and natural in some way. Anthony Smith (1986) identifies a variety of different kinds of primordialisms, including *essentialist primordialism* and *kinship primordialism*. Essentialist primordialism holds that ethnicity is a natural biological fact, and thus ethnicity precedes human society. This approach has had relatively little impact on scholarship, although it has been forwarded politically by racial supremacists from time to time. More common in the primordialist literature is kinship primordialism, which holds that ethnic groups are the extensions of blood kinship communities. A version of this is the perspective offered by Clifford Geertz, who acknowledges that ethnic identity is not entirely “in the blood” but that ethnic ties and group bonding are somewhat natural processes (because objective racial or physical features assist the creation of social bonds). In particular, from this point of view, ethnic ties represent a permanent social bond that is self-sustaining and not subject to human manipulation (Geertz, 1973). Indeed, Geertz argues that because people attach so much to certain socially objective markers such as race, religion, language, and culture, they perceive ethnicity as primordial and natural.

On the other hand, the *situational perspective* (also known as the *constructionist* or *instrumentalist* approach) states that ethnic identities are socially constructed. In other words, the definition of the group, and the identification of its boundaries, are often negotiated and renegotiated, and how these boundaries are redefined depends on specific situations and circumstances that each group encounters. The basic cognitive processes of self-categorization and self-schematization (Turner, 1985), combined with social interaction (Burke & Reitzes, 1981), produce intersubjective agreement that (almost) every person can be placed into categories described in terms of some easily perceived attributes. People attach themselves to these groups, and the strength of these attachments or commitments (Burke & Reitzes, 1981) affects people’s lives and activities.

The instrumentalist approach is more cynical than the simple notion that identities are social constructions. Instrumentalists see the creation of identity as the product of the manipulation of cultural and kinship symbols by political entrepreneurs for political gain (Cohen, 1974). This approach sees ethnicity as result of political strategy, usually to achieve other ends, such as political

power, access to resources, and increase in wealth and status.

Within the situational perspective, several subtheories attempt to explain how ethnic identity is formed and reshaped. Some sociologists, for example, have argued that ethnic identity can be resurgent or emergent. Those who believe in a *resurgent* ethnic identity accept the idea that traditional or ancestral identities can reemerge as the result of particular historical circumstances. For instance, groups that have been stripped of their previous heritage (such as Native Americans) or assimilated (such as many groups in the former Soviet Union) have experienced a resurgence of pride in traditional identities. This has happened in the 20th century in North America and, in the former Soviet Union after the collapse of Communism, for ethnic minority groups such as the Tatar and Bashkir and new majority groups such as Kazakhs and Kirghiz.

On the other hand, an emergent ethnic identity involves the creation of a new sense of group identity that may emerge as the result of particular circumstances. For example, Japanese Americans in the United States, largely as the result of World War II, have formed an identity quite different from being either Japanese or entirely “American” (Nagel, 1994). Indeed, as policies and context lead to contested identities, individuals may go so far as to create new identities, which become shared.

Last, considerable recent interest has focused on how ethnic identities change, particularly after group circumstances change. This approach is based on the notion that contextual changes affect the relationships between peoples. As several scholars have noted, economic, social, and political changes shape the context affecting groups and individuals (Anderson, 1983; Hobsbawm, 1990; Posner, 2005). This is because the very attributes that shaped identity before are suddenly in flux. During times of disruption (e.g., in post-Soviet politics), dimensions such as language, tribe, or region may be used by political entrepreneurs to restructure groups based on new identities (Posner, 2005).

With the emergence of newly independent nations from the former Soviet Union and the former Yugoslavia, the issue of ethnic and national identity has changed as well. Changes in the ranking of individual groups affect the development of identities. The character of the ranking systems (Horowitz, 1985) affects whether individual ethnics will try to change identities or must attempt to change the status of their group in order to improve their conditions. Horowitz (1985) contrasts the two systems as follows:

[In unranked systems] one need not choose between his mobility aspirations and his group membership, whereas in ranked systems elite status is possible for members of a subordinate group only if they are willing and able to renounce their origins by passing into the superordinate group. (p. 35)

Thus, for example, groups such as Kazakhs in Kazakhstan, which were ranked below Russians in Kazakhstan during the Soviet period, are now ranked above Russians in Kazakhstan. This makes Kazakh identity more attractive than a Russian identity for many who live in Kazakhstan.

What Is the Relationship Between Ethnic and National Identity?

A second debate in the literature on ethnic identity, as identified by Thomas H. Eriksen (2001), is the debate within nationalist studies between constructivism and essentialism, or how national identities emerge as a product of ethnic identities. This involves the debate over whether national communities are created consciously or whether they grow organically out of preexisting cultural communities. In nationalism studies, this debate is highlighted by the differences between Smith (1986) and Ernest Gellner (1983). Gellner argues that nations are entirely a product of modernization. Gellner is not alone in this point of view. For instance, the Marxist historian Eric Hobsbawm also argues that ethnicity and nationalism are wholly modern inventions, appearing only after industrialization and modernization in world history. Thus, from this point of view, industrialization and urbanization created the basis for mass national identities. For instance, prior to industrialization, national identities did not exist. Generally, identities were local, not national. Thus, for instance, a Romanian peasant in the early 19th century was likely to identify with his or her local village or district, not with a community known as Romania, which did not exist politically (the region now known as Romania was then part of the Ottoman Empire). However, as economic modernization advanced, accompanied by the movement to urban areas, people became connected in ways that transcended local identities. People realized that there were now others who spoke the same language and who shared the same cultural practices. This created the basis for a mass identity, such as Romanian. Subsequently, the state creates a new history that proclaims that the nation has ancient roots that have existed for millennia (Gellner, 1983).

Benedict Anderson (1983) also argues that in many ways nations are “imagined” communities that are created, although unlike Gellner, he sees no necessary connection with particular ethnic groups. Gellner sees nations growing out of ethnic communities (such as the Romanian example above). Anderson, on the other hand, observes that nations can include many ethnic groups but that the national community is a community that binds people together nonetheless (for Anderson the Philippines and Indonesia are cases in point).

On the other hand Smith (1986, 1991) has argued that every nation has an ethnic core and that nations are not

simply imagined. He has pointed to the importance of preexisting *ethnies*, which are required for the development of nationalism. An *ethnie* is defined as a population with a shared proper name, a common ancestry, shared historical memories, a shared culture, and an association with a particular territory. Nations grow out of these groups. Indeed, Smith argues that nationalism draws on the history of particular groups to fashion a sense of identity.

Greenfeld (1992) shares this perspective and argues that nations are not just “imagined communities.” However, she says that national identities are not completely fixed, either. Rather, whether an individual chooses to identify with a particular national community depends to a large extent on the individual’s personal experiences. In *Nationalism: Five Roads to Modernity*, she outlines the process by which individuals link their personal experiences with the development of a collective identity or nation. Greenfeld contends that one must experience a social, political, psychological, religious, or cultural event that encourages the person to accept a community. As rites and initiations occur, the meanings of these situations build on each other until the individual identifies within the community and accepts or rejects its traditions. In either case, the person, through self-selection or marriage, connects with a group that creates a community and a nation. Further, several social institutions support the sustenance of a community: family, schools, peers, religion, the community norms, its media, and life events.

A related issue regarding the creation of national identity is the question of who belongs to the national community—that is, what are the criteria for membership in the national club. Indeed, several scholars have distinguished between communities that have different criteria for membership, particularly the difference between *ethnic* nations and *civic* ones (Smith, 1991). Rogers Brubaker (1992), for instance, found that the French and German conceptions of nation and nation-state are very different. France has historically emphasized loyalty to the state and acceptance of French culture and ideas as the key criteria for citizenship. Thus there is an emphasis on assimilation of peoples, regardless (technically) of race or other physical characteristics. On the other hand, the German conception of national community is based on the concept of the people (*Volk*), or the idea that an ethnocultural community is based on “blood.” These divergent approaches to national community partially explain the different ways in which countries approach issues of citizenship and immigration. In France, citizenship is based on acceptance of the French language and culture. Immigrants are expected to assimilate into the French community by adopting these attributes. In Germany, on the other hand, the emphasis on an ethnic community means that blood connections are far more important as a criterion for membership in a national community than is cultural assimilation. In fact, cultural

assimilation is simply not enough to gain membership in the national community.

Greenfeld (1992) argues that these different kinds of national identities affect the kinds of nationalisms that emerge. She suggests that there are three basic types of nationalism: individualist-civic, collectivist-civic, and collectivist-ethnic. *Individualist civic* nationalism identifies a collection of individuals who come together in civic life and abide by its conventions. The national community is primarily held together by some loyalty to a set of principles, as opposed to some mystical notion of *volk*. Nationalism in the United States represents this type. In *collectivist civic* nationalism, loyalty to the state defines membership in the national community, and this loyalty is often accompanied by a demand to assimilate to the dominant culture. France represents this type of nationalism. Finally, there is *collectivist ethnic* nationalism, which is the most exclusionary and is based on the idea that the national community is defined by blood. Germany, particularly in the past, would fit this type of nationalism. Often this form of nationalism is directed from above and exclusionary of outsiders.

Some Research Directions

The issue of ethnic identity has been an important research issue in comparative political science. Two research agendas illustrate some of the scholarship in the area: (1) the effects of ethnic identity on voting and (2) the effects of ethnic parties on new democracies.

Ethnic Voting

Scholarship has a long-standing interest in elections and voting behavior in “divided” societies. Indeed, many scholars have worried that elections in ethnically divided societies will produce *census elections*, which are inimical to democracy. In other words, such elections tend to create impermeable blocs that detract from interethnic accommodation. Rather than creating harmony and stability, elections in ethnically divided societies become an invitation to fraud and open conflict as groups struggle for political dominance (Horowitz, 1985; Rabushka & Shepsle, 1972). The most common explanation for census-style elections is Horowitz’s (1985) *expressive voting hypothesis*, which contends that voters employ the act of voting to express (and hence register) their identities as part of an ethnic group. This practice in turn gives rise to ethnic parties, which, through the process of *ethnic outbidding*, lead to the hardening of ethnic positions, which reduces the possibility of ethnic accommodation and ultimately leads to the dissolution of incipient democracies.

A prominent theme among scholars studying voting in new democracies is the enduring relevance of ethnicity in

the developing world, where “ethnic ties based on kinship and family, language and dialect, tribal customs and local communities, as well as shared religious faiths, have long been regarded as playing a critical role in party politics” (Norris & Mattes, 2003, p. 1). For some, such as Letitia Lawson (1999), ethnicity is an alternative basis for political mobilization and an almost natural process in Asia and Africa, a process she describes as follows:

Absence of formal associations clearly apart from the state and capable of engaging the population, the introduction of liberal democratic procedures, at the behest of external donors, [all this] has led political parties to appeal to the only available alternative: ethnic identity. (p. 12)

While there is a general consensus that ethnicity is an important cleavage in many countries, there has been considerable debate over *why* voters may vote along ethnic lines. Some have argued that voters vote for ethnic candidates because they believe that “their” candidates will deliver patronage resources to them. For instance, Chandra argues that ethnic voting is often the product of *patronage democracies*, in which the state monopolizes access to jobs and administrative resources (Chandra, 2004). Such states, in the search for legitimation and popular support, often engage in *selective targeting*, whereby certain groups are selected to receive the most benefits from state resources. Hence, voting aligns along ethnic lines largely out of the expectation for patronage benefits. In this way, ethnic identity offers an informational shorthand, telling voters that voting for a coethnic candidate or party will more likely result in benefits for the voter than will voting for a candidate or party from another ethnic group. Thus, voters believe that benefits are more likely to result from demonstrating solidarity with the community. In other words, they share a sense “that only a member of their own ethnic group may end up defending the interests of the ethnic group as a whole, and that voting for a member of another ethnic group will certainly not do so” (van de Walle, 2003, p. 313).

A second approach emphasizes the coincidence of ethnicity with policy preferences and voter evaluations of the performance of the government. Indeed, it may be the case that what appears to be an ethnic bond among voters is actually the result of voters’ sharing common policy preferences that drive them to vote in similar ways. For instance, Robert Mattes (1995) notes that in divided societies, individuals from the same ethnic group have common preferences because they often share common political and economic interests. Because preferences form as the result of social interactions and individuals within ethnic communities tend to interact only with other members of their own group, shared patterns of interests result. Thus, it may appear that ethnicity is the cause of voting behavior, but in reality ethnicity is merely a correlate of other interests. Karen Ferree (2004) refers to this as

the *policy framework approach*, whereby voting is driven not by identity but by a common set of interests. Indeed, voter choice often appears to be primarily reflective of ethnic identity when it is actually reflective of a common set of frustrations with government performance. In this situation, ethnicity plays no direct role in shaping voting behavior. That ethnicity overlaps with behavior is largely happenstance.

Last, the *ethnic expressive voting* approach suggests that voters vote for coethnic candidates because the act of voting is a psychological affirmation of group identity. The most prominent advocate of this approach is Horowitz (1985). Horowitz argues that to vote for an ethnic party is to affirm group identity, and voters thus derive psychological benefits from supporting ethnic parties. Since voting is not a product of rational calculation, partisan allegiances are rather fixed and rigid, and elections become a reflection of demographics. Such voting implies that voting is not the outcome of a careful evaluation of policy positions or the performance of leaders; instead, it is identity that matters. As Stanley Fish (2008) notes, the heart of identity politics is that voters vote for or against someone because of the candidate’s personal characteristics. “In essence, identity politics is an affirmation of tribe against the claims of ideology” (p. A13).

Ethnic Parties, Conflict, and Accommodation

Related to ethnic voting is the topic of ethnic parties, which is often seen as the organization product of ethnic identity voting. Indeed, there has been a long-standing interest in the role played by ethnic parties in promoting (or dampening) conflict between ethnic groups. Many scholars have argued that the appearance of ethnic parties is a “bad” thing for new democracies or systems in transition. From this perspective, not only does the emergence of ethnic parties deepen divisions between groups, but ethnic parties serve to exacerbate conflict (Hislope, 1997).

Alvin Rabushka and Kenneth Shepsle (1972) offered one of the earliest explanations of the way in which ethnic parties promote conflict in plural societies. In their model, a key role is played by ethnic elites and the organizations they lead. Indeed, these organizations engage in the politics of ethnic outbidding that ultimately undermines multi-ethnic cooperation and inevitably leads to nondemocratic, ethnically exclusive states. Other authors have similarly argued that because ethnic parties make their political appeal specifically on ethnicity, their emergence often has a centrifugal effect on politics. This effect is especially harmful to new democracies, where democratic institutions are quite fragile. Indeed, under such conditions, ethnic competition can easily turn into ethnic conflict. This is because the competition for votes for the ethnic party involves mobilizing the ethnic group, and the best way to do that is to use inflammatory and confrontational rhetoric, distinguishing between “us” and “them” (Koelble, 1995).

Richard Gunther and Larry Diamond (2003) have described this process as follows:

The electoral logic of the ethnic party is to harden and mobilize its ethnic base with exclusive, often polarizing appeals to ethnic group opportunity and threat. . . . The ethnic party's particularistic, exclusivist, and often polarizing political appeals make its overall contribution to society divisive and even disintegrative. (pp. 21–23)

In addition, ethnic parties indirectly contribute to worsening ethnic tensions by promoting party politics along cultural lines, which often leads to the marginalization and exclusion of a cultural minority. As a consequence, such minorities may feel encouraged to resort to undemocratic or even violent means in order to counter this treatment. Further, merely by promoting identity-based politics, ethnic parties can significantly raise the stakes of the political game, reinforcing group identities and thus raising the likelihood of conflict. For these reasons as well, ethnic parties increase the likelihood of intercommunal conflict and threaten the survivability of new democracies.

Thus, from the foregoing perspective, the mere appearance of an ethnic party should signal an increase in the rise of interethnic conflict. Horowitz (1985) has described the phenomenon in the following terms:

By appealing to electorates in ethnic terms, by making ethnic demands on government, and by bolstering the influence of ethnically chauvinistic elements within each group, parties that begin by merely mirroring ethnic divisions help to deepen and extend them. Hence the oft heard remark in such states that politicians have created ethnic conflict. (p. 291)

On the other hand, several scholars contend that ethnic parties can play a constructive role in promoting inter-group accommodation. Indeed, advocates of the consociational school have long argued that promoting the emergence of ethnic parties and then representing them broadly will facilitate the integration of as many subcultures as possible into the political game, thus creating the conditions for interethnic cooperation (Daalder, 1974; Lijphart, 1968). Furthermore, securing representation for minority groups facilitates the integration of disaffected groups into the political system, which ultimately leads them to moderate their demands. Frank Cohen (1997) argues that the broader the representation, the more likely the ethnic group feels bound to the existing system. As Cohen puts it, "By making institutions more accessible and making ethnic cleavages more explicit, ethnic groups will engage in more frequent but less intense conflict. They will use moderate means of resistance to effect change in the status quo" (p. 613).

Others, such as Sherrill Stroschein (2001), contend that ethnic parties do not cause ethnic conflict but emerge as the result of it. That is, they reflect differences that already

exist. Nonetheless, ethnic parties can channel demands into more legitimate forms of participation and thus allow conflicts to be resolved politically rather than through violence (p. 61). John T. Ishiyama (2000) demonstrated that in the post-Communist world, ethnic parties have assisted in bringing into the political process those who otherwise would have been alienated by the emerging democratic systems in the region.

Current research, such as that by Johanna Birnir (2007), Chandra (2004), James Fearon and David Laitin (1996), and Daniel Posner (2004), offers a more contingent view of the link between ethnic cleavages and conflict and stresses the importance of ethnicity and ethnic cleavages as cost-effective strategic resources for group formation, interest definition, and collective action. Perhaps one of the strongest and most articulate proponents of the notion that ethnic parties can have a positive effect on the stabilization of new democracies is Chandra (2005). She directly criticizes the notion of ethnic outbidding, which is so central to the argument that the mere appearance of ethnic parties sets off a chain reaction leading to a spiral of extremism that destroys democratic politics altogether. Rather, she argues that ethnic parties can help sustain democracy if these parties are institutionally encouraged to compete on multiple dimensions rather than on just the unidimensional axis of ethnicity. Indeed, political institutions that restrict "ethnic politics to a single dimension destabilize democracy, whereas institutions that foster multiple dimensions of ethnic identity can sustain it" (Chandra, 2005, p. 236).

In a similar vein, Birnir (2007), in examining patterns of ethnic politics in a broadly comparative way, contends that the ethnification of politics does not necessarily translate to violence. Indeed, like Chandra, Birnir argues that ethnic identity serves as a stable but flexible information shorthand for political choices and assists in stabilizing party formations and hence the development of democracy. If violence results, it is largely the result of political institutional factors, particularly restrictions on access to the executive. This exclusion is what leads to violence, not the political mobilization of ethnicity.

Birnir argues that, *ceteris paribus*, ethnic parties (which she refers to as ethnic *attractors*) are predisposed to seek peaceful means to gain access to political power. This is because, as with all parties, ethnic attractors seek to act on behalf of a constituency and seek leverage for that constituency. In turn, voters who use ethnic identity as a shortcut to sort through candidate preferences prefer parties that act on behalf of the ethnic constituency (this party could be a nonethnic party as well). This factor provides a strong incentive for the ethnic attractor to gain access to the political executive, which is best achieved through peaceful means.

Why is it the case, then, that members of some ethnic groups appear to peacefully support their group in electoral

politics whereas others do not support their groups, exit electoral politics, and even engage in protest and violence? Birnir's answer is that if political intransigence and violence result, it is not because of the ethnification of politics but rather the denial of political access to an ethnic group. It is the shutting out from the core of power that produces the kinds of violence and instability that are commonly associated with ethnic politics in the existing literature.

Conclusion

This chapter has sought to summarize the basic literature on ethnic identity politics, a subfield of comparative politics. Although its coverage is not entirely exhaustive, it provides the reader with a basic understanding of the underlying debates regarding the central questions addressed in the literature: (a) What is ethnic identity? (b) What leads to the formation of an ethnic identity? (c) What is the relationship between ethnic identity and political movements such as nationalism? (d) How are issues of ethnic identity explored in the literature on comparative ethnic voting and ethnic parties? Furthermore, with the changes wrought by the end of the cold war and the emergence of new states and new identities from the former Soviet Union, as well as in other parts of the world, the answers to questions about how identities change are now accessible to scholars. Indeed, unlike in the past, when analyses of national and ethnic identity formation were largely derived from secondhand observations of history, the process of identity formation is unfolding in front of us. This provides new opportunities and new research directions that represent exciting new paths for future inquiry.

References and Further Readings

- Anderson, B. (1983). *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.
- Barth, F. (Ed.). (1969). *Ethnic groups and boundaries*. Oslo, Norway: Scandinavian University Press.
- Birnir, J. (2007). *Ethnicity and electoral politics*. New York: Cambridge University Press.
- Brubaker, R. (1992). Civic and ethnic nations in France and Germany. In J. Hutchinson & A. Smith (Eds.), *Ethnicity* (pp. 168–173). Oxford, UK: Oxford University Press.
- Burke, P., & Reitzes, D. C. (1981). The link between identity theory and role performance. *Social Psychology Quarterly*, 44, 83–92.
- Chandra, K. (2004). *Why ethnic parties succeed: Patronage and ethnic head counts in India*. New York: Cambridge University Press.
- Chandra, K. (2005). Ethnic parties and democratic stability. *Perspectives on Politics*, 3(3), 235–252.
- Chandra, K. (2006). What is ethnic identity and does it matter? *Annual Review of Political Science*, 9, 397–424.
- Cohen, A. (1974). *Two dimensional man: An essay on power and symbolism in complex society*. London: Routledge & Kegan Paul.
- Cohen, F. (1997). Proportional versus majoritarian ethnic conflict management in democracies. *Comparative Political Studies*, 30(5), 607–630.
- Cohen, R. (1978). Ethnicity: Problem and focus in anthropology. In B. J. Siegel (Ed.), *Annual review of anthropology* (Vol. 7, pp. 379–403). Palo Alto, CA: Stanford University Press.
- Daalder, H. (1974). The consociational democracy theme. *World Politics*, 26(4), 604–621.
- Eriksen, T. H. (2001). Ethnic identity, national identity and intergroup conflict: The significance of personal experiences. In R. D. Ashmore, L. J. Jussim, & D. Wilder (Eds.), *Social identity, intergroup conflict, and conflict reduction* (pp. 42–70). Oxford, UK: Oxford University Press.
- Fearon, J., & Laitin, D. (1996). Explaining interethnic cooperation. *American Political Science Review*, 90(4), 715–735.
- Fearon, J., & Laitin, D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1), 75–90.
- Ferree, K. (2004). *The microfoundations of ethnic voting: Evidence from South Africa* (Afrobarometer Working Paper No. 40). Retrieved May 30, 2010, from <http://www.afrobarometer.org/abseries.html>
- Ferree, K. (2006). Explaining South Africa's racial census. *Journal of Politics*, 68(4), 803–815.
- Fish, S. (2008, February 17). When "identity politics" is rational. *The New York Times*. Retrieved May 10, 2010, from <http://opinionator.blogs.nytimes.com/2008/02/17/when-identity-politics-is-rational/>
- Geertz, C. (Ed.). (1967). *Old societies and new states: The quest for modernity in Africa and Asia*. New York: Free Press.
- Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.
- Gellner, E. (1983). *Nations and nationalism*. Ithaca, NY: Cornell University Press.
- Greenfeld, L. (1992). *Nationalism: Five roads to modernity*. Cambridge, MA: Harvard University Press.
- Gunther, R., & Diamond, L. (2003). Types and functions of parties. In L. Diamond & R. Gunther (Eds.), *Political parties and democracy* (pp. 3–40). Baltimore: Johns Hopkins University Press.
- Hislop, R. (1997). Intraethnic conflict in Croatia and Serbia: Flanking and the consequences for democracy. *East European Quarterly*, 30(4), 471–494.
- Hobsbawm, E. J. (1990). *Nations and nationalism since 1780*. New York: Cambridge University Press.
- Horowitz, D. (1985). *Ethnic groups in conflict*. Berkeley: University of California Press.
- Ishiyama, J. (2000). Institutions and ethno-political conflict in post Communist politics. *Nationalism and Ethnic Politics*, 6(1), 51–67.
- Koelble, T. (1995). Towards a theory of nationalism: Culture, structure and choice analyses revisited. *Nationalism & Ethnic Politics*, 1(4), 73–90.
- Lawson, L. (1999). External democracy promotion in Africa: Another false start? *Commonwealth and Comparative Politics*, 37(1), 1–30.
- Lijphart, A. (1968). *The politics of accommodation: Pluralism and democracy in the Netherlands* (2nd ed.). Berkeley: University of California Press.

- Mattes, R. (1995). *The election book: Judgment and choice in South Africa's 1994 election*. Cape Town, South Africa: Idasa.
- Nagel, J. (1994). Constructing ethnicity: Creating and recreating ethnic identity and culture. *Social Problems*, 41(1), 152-176.
- Norris, P., & Mattes, R. (2003). *Does ethnicity determine support for the governing party?* (Afrobarometer Working Paper No. 26). Retrieved May 30, 2010, from <http://www.afrobarometer.org/abseries.html>
- Posner, D. (2004). Measuring ethnic fractionalization in Africa. *American Journal of Political Science*, 48(4), 849-863.
- Posner, D. (2005). *Institutions and ethnic politics*. New York: Cambridge University Press.
- Rabushka, A., & Shepsle, K. (1972). *Politics in plural societies: A theory of democratic instability*. Columbus, OH: Merrill.
- Smith, A. (1991). *National identity*. London: Penguin.
- Smith, A. D. (1986). *The ethnic origins of nations*. Oxford, UK: Blackwell.
- Stroschein, S. (2001). Measuring ethnic party success in Romania, Slovakia and Ukraine. *Problems of Post Communism*, 4(1), 59-69.
- Stryker, S., & Serpe, R. R. (1994). Identity salience and psychological centrality: Equivalent, overlapping, or complementary concepts? *Social Psychology Quarterly*, 57, 16-35.
- Turner, J. C. (1985). Social categorization and the self concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2., pp. 77-121). New York: JAI Press.
- van de Walle, N. (2003). Presidentialism and clientism in Africa's emerging party systems. *Journal of Modern African Studies*, 41(2), 297-322.
- Vincent, J. (1974). The structure of ethnicity. *Human Organization*, 33(4), 375-379.

27

SOCIAL MOVEMENTS

MARC HOOGHE

University of Leuven

Social movements can be conceptualized as sustained and enduring challenges to political decision makers in order to achieve some form of social change. Although social movements most often are composed of one or various social movement organizations, various authors have emphasized that social movements should not be identified solely with those organizations. Individual actions, cultural manifestations, the activity of opinion leaders and other elements of cultural change, and consciousness-raising can also be labeled as elements of social movements. Although social movements are studied mostly within the field of sociology, they are also of crucial importance within political science. It can be argued that some of the most important political changes in the 19th and 20th centuries were brought about by the actions of social movements. Powerful examples are the civil rights movements in the United States, the green movement, and women's organizations, but one could also think about organizations aimed at promoting gay rights or the protest against authoritarian regimes in Central and Eastern Europe in the 1980s. Social movements therefore are usually identified with contentious politics: They try to bring about political change by challenging the political elite. As such, they give voice to those who have been excluded from the political system.

Social movements can be distinguished from political parties as they do not directly compete in elections and they do not participate in exercising state power by means

of elected officials. In various instances, however, there are examples of social movements that are linked quite narrowly to political parties. Social movements can also function as interest groups, but mass participation of citizens is usually considered as a defining element of social movements, and it is not necessarily a defining element of interest groups. This distinction from political parties does not prevent social movement organizations from trying to have a direct impact on political decision-making processes. Participation is thus an important element of social movements, and although social movement participation cannot be considered as a standard form of institutionalized political participation, most authors agree on the fact that social movement participation is or has become an important element in the political action repertoire of citizens in liberal democracies. Social movement organizations themselves use a wide array of activities (demonstrations, strikes, newsletters, lobbying techniques, print, etc.) to reach their goals of mobilizing their constituency and to have an impact on political decision making.

Within political science, the study of social movements has received considerable attention. Whereas in the first half of the 20th century, social movements were usually depicted as a potential threat to political stability, the idea emerged, in the 1960s especially, that social movement participation could be seen as a normal and even necessary element within democratic political systems. In the development of the study of social movements,

scholars have successively emphasized the resources that are necessary to sustain social movement organizations, the cultural meaning of social movements, and the political conditions that facilitate the occurrence and the success of social movements. In the more recent literature, authors have highlighted the network structure of movements, the emotional motivations of participants, or the political and cultural consequences of social movement activity. From a review of the current literature, it seems clear that there is no longer one dominant paradigm in the study of social movements. Rather, it is acknowledged that in order to arrive at a comprehensive study of the phenomenon, organizational, cultural, and political insights should be combined.

Development of Social Movement Studies

Early Approaches

In the early 20th century, social movements were mostly studied as a form of collective behavior, and this line of research was heavily influenced by the insights of crowd psychology. It was feared that individuals would lose their rationality once they participated in crowds because they would give in to a propensity to follow crowd leaders. Collective behavior could easily lead to mass gatherings, mobs, riots, and even forms of collective violence. This kind of crowd behavior was considered a potential threat to the stability of the political system. This negative conception of social movements can be explained by two distinct elements. On one hand, it should be remembered that this era indeed witnessed a number of unruly forms of participation, leading to political violence, attacks, and other forms of social disruption. On the other hand, it was also clear that sociologists and political scientists apparently identified with the current political and social order, seeing attempts to change the structure of society as a threat to political stability. In other scientific disciplines, too, an elite distrust of mass behavior was quite clearly present during this period. This conservative outlook was especially present in the seminal work of the French social psychologist Gustave Le Bon (1841–1931). In his work published in English under the title *The Crowd* (1895/1977), Le Bon stated that participants in mass gatherings displayed a tendency to behave in a herdlike manner, blindly following their leaders. The stability of society was ensured by a rational elite that was able to oppose the claims of unruly mobs.

This negative outlook toward mass participation was questioned strongly in the 1959 volume *The Politics of Mass Society*, by Berkeley sociologist William Kornhauser (1959/2008). Kornhauser radically reversed the Le Bon framework. He argued that individual and passive citizens normally feel isolated from the social order because they are powerless to bring about social change on their own. Their

participation in social movements, on the other hand, leads to a feeling of empowerment because it allows them to reach collective goals and to bring about social change. As such, social movements actually contribute to the social order by providing a mechanism for social innovation and social integration. Without social movements, citizens feel isolated and alienated from the social order. While some social movements might aim to overthrow the social order, other movements are aimed at implementing reforms that could ensure the long-term stability of society and bring the functioning of society in line with its professed value preferences.

The civil rights movement and the student uprisings of the 1960s, however, clearly forced political scientists to reconsider their view on social movements in a more profound manner as it became clear that these movements did not recruit just alienated members of society, and it became equally clear that these social movements would have profound political and social consequences. For the first time, academics themselves started to participate actively in these kinds of movements. It has to be noted in this respect that the writings of social historians such as Charles Tilly, E. P. Thompson, and E. J. Hobsbawm can be considered as pioneering contributions to this line of study, with an influence that was present in various scientific disciplines. In 1964, for example, Tilly published his seminal study on the counterrevolutionary uprising in the Vendée in 1793, explaining this movement by stressing the economic and structural strains the population of that region of France experienced. Other historians also arrived at the conclusion that uprisings no longer should be seen as just an emotional reaction, devoid of historical or political meaning, but that they actually could be interpreted as the result of a number of structural social changes. English historians stressed the moral indignation over economic injustice that gave rise to protest behavior in the 18th century. The writings of these social historians also had a profound influence on the way political scientists came to regard social movements.

Resources

A first and crucial innovation was the introduction of the *resource mobilization theory* by sociologists John McCarthy and Mayer Zald (1977). They argued that social movements are not just a spontaneous reaction to grievances and forms of discontent. Like all other forms of collective behavior, social movements are also dependent on a supply of material resources, such as time, money, pre-existing organizational structures, or organizational skills. This way of looking at social movements was heavily dependent on rational choice approaches to human behavior. The underlying idea was that participants in social movements do not take part just out of some frustration or discontent but that they will embark on participation only if they arrive at the conclusion that this investment (in time or energy or risk-taking behavior) can be considered worthwhile or will lead to some form of social or political

change. If organizational resources are not present at all, it does not make much sense to take part in contentious behavior. During the 1970s and 1980s, this resource mobilization theory was very influential in the study of social movements as it allowed scholars to investigate movements as a rational form of collective action. It also provided them with a checklist of items that were considered essential for the occurrence of forms of collective action. Organizational structures, means of communication, and preexisting recruitment networks were all considered necessary to mount a successful social movement organization. Critics, however, argued that this approach neglected the specific character of social movements, as the same rational approach could be applied to any form of political behavior. They also argued that various movements succeeded in bringing about social and political change in the absence of material resources.

Cultural Approaches

While the resource mobilization theory was very influential in the United States, it never became the dominant approach in European studies on social movements. Like their U.S. counterparts, European political scientists were surprised by the rise of the new social movements of the 1960s and 1970s, and they had to scramble to explain the phenomenon in a coherent manner. Most of them resorted to some form of *strain theory*: it was argued that rapid social change imposed strains on the functioning of society, and social movements were seen not just as a manifestation of those strains but also as a viable way to arrive at solutions for this kind of social pressure. Rapid industrial and population growth, for example, led to the occurrence of strains on the stability of natural ecosystems, and the newly emerging ecological movement not only focused attention on those strains but also developed possible solutions to arrive at a new environmental equilibrium. Social movements often were credited with playing a pioneering role because they succeeded in putting new challenges and new items on the political agenda. Within this approach, the term *new social movements* was used, in order to make a distinction between the classic social movements of the 19th century (e.g., trade unions) and the new social movements of the late 20th century (the women's movement, civil rights, ecology, etc.). New social movements were seen as a very typical phenomenon related to a specific developmental phase of highly industrialized societies. That is, the new social movements marked the transition from industrial to postindustrial societies.

The problem with this approach, however, was that it failed to explain the continuities in various forms of collective political action. Microstudies showed that the challenges (with regard to recruitment, mobilization, building political influence, etc.) facing new social movements were not all that different from the challenges that the social movements of the 19th century faced. Research also revealed continuities, such as between the women's suffrage movement of the

early 20th century and the feminist movements of the 1960s or between the nature conservation movement of the late 19th century and the modern ecological movement. Furthermore, the direct causal link between structural social change and the occurrence of new social movements failed to explain why some social movements were more successful in some countries than in others. If, for example, the green movement should be seen as a direct reaction to the degradation of the natural environment, one would expect this movement to be present—and successful in mobilizing—in all highly industrialized societies experiencing environmental degradation. This, however, was clearly not the case. Since the 1990s, in fact, quite a few of these new social movement organizations had entered into a phase of demobilization, so it became increasingly difficult for this line of research to attribute new social movements with a structural role in processes of social change.

Political Opportunities

One of the crucial questions in the comparative study of social movements is why these movements are successful in some countries and not in others, given the fact that grievances can be considered as universal. A seminal study by Doug McAdam (1982) started from the observation that the American civil rights movement started in the 1950s, at a moment when racial discrimination in American society actually started to decline. According to McAdam, the civil rights movement could prosper as a result of growing divisions within the ruling White elites. This meant that the civil rights movement had the opportunity to exploit those divisions as it gained access to elite allies that were instrumental in furthering the policy goals of the movement. The important lesson from this study is that social movements should not be seen as a direct reaction to grievances or specific problems within society, but that they can be seen as reactions to changes in the political system. Social movement organizations respond to the opportunities created by the system.

This *political opportunity structure* approach was later systematically ordered mainly in the work of the Swiss political scientist Hanspeter Kriesi and his research team (Kriesi, Koopmans, Duyvendak, & Giugni, 1995). The main idea of this approach is that the chances for success and mobilization of social movement organizations are strongly dependent on the opportunities created and offered by the political system. These opportunities can be institutionalized and formal (e.g., freedom of assembly, openness of procedures), but they can also be informal (e.g., cultures of pluralism or corporatism). Authorities or political elites can also try to facilitate or to repress social movement manifestations, and this will have an impact on a movement's chances for success. The political opportunity structure approach was successfully used in a number of comparative studies, demonstrating that the success of social movements in some countries was strongly dependent on features of the political systems in those countries.

Critics of the political opportunity structure approach argued that this way of looking at social movements limited the agency of these movements, as they were sometimes seen as a simple reaction to decisions being made by the political system. Proponents of the approach responded by introducing feedback loops into the model: Self-evidently, social movements can influence even basic characteristics of the political system (such as universal voting rights), so that social movement organizations in a later developmental phase will profit from the precedents created by earlier organizations.

Comprehensive Overview

What these three lines of theory building on social movements have in common is that they depart from a positive outlook on social movements and their effects on the political system. Participation in movements is seen as a rational act and a form of participation that is congruent with the prevalence of a democratic political culture. Other authors would even argue that social movements are necessary in order to bring about social change and a further democratization of contemporary societies. As a form of criticism, however, it has also been argued that each one of these theoretical perspectives offers a one-sided outlook on social movements. Self-evidently, material resources are important, but it is only if they can be used in an open political context, and if movements have access to a political and social culture that is conducive to their goals, that they will have an opportunity to succeed and to prosper and to have a profound impact on the political decision-making process.

An impressive effort to synthesize these approaches was published in 1996, when McAdam, McCarthy, and Zald tried to combine the insights from this research into political opportunities, mobilizing structures, and cultural frames. The guiding idea of this volume was that the various approaches had to be combined if one wanted to study the development of social movements in a comprehensive manner. Although this study has been very influential, one cannot state that it has led to a unification of the theoretical perspectives on social movements. Following all the elements that were brought forward in these approaches would amount to constructing a very long checklist of all conditions that are conducive to the success of social movement organizations, and this apparently was judged as not very practical. In reality, therefore, most of the researchers in this field continue to focus on either the material and organizational, or cultural, or political aspects of the functioning of social movements.

Empirical Research

It is difficult to develop a comprehensive account of currently available empirical research on social movements. Some studies focus on particular movements; others opt

for a comparative approach (across issues or across countries); still others focus on issues of recruitment, organization, framing of issues, or political consequences (Walder, 2009). This section focuses on a number of recent influential studies that take a lifestyle approach, starting from the origins of social movements to their political impact.

In the 1960s and 1970s, some authors were still looking for an objective assessment of grievances and social demands that could lead to the foundation of social movements. It was claimed that a sudden deterioration of living conditions (e.g., with regard to the environment, gender, or racial discrimination), or on the contrary a gradual improvement in experienced living conditions, would give rise to social movements. An important insight was that it was not the actual quality of life that had an impact on grievances but rather the difference between expected and perceived quality of life. In the 1980s, this *grievances approach* was gradually replaced by the *framing approach* to social movement demands (Benford & Snow, 2000). The guiding idea of this line of research is that social movements make claims about society and about ways that society should and can be changed. Collective action frames are successful if they succeed in mobilizing a sufficient number of participants and if they convey the message that change is indeed possible. Frames are more likely to be successful if they resonate with preexisting concepts within public opinion about how society is managed or should be managed. Studying claim-making activities as a form of framing reality implied in practice that researchers no longer tried to link the occurrence of social movements to real-life conditions.

Cultural frames, however, are not sufficient to start a successful social movement. One also needs participants. Research has shown repeatedly that individual participants are not recruited in an organizational void. On the contrary, participants are mainly recruited in preexisting networks or in various contexts (schools, work environments, neighborhoods, etc.) that are conducive to recruitment. This form of group recruitment can be explained partly by cultural mechanisms as it is likely that citizens who are active in the same kind of context will share at least some political preferences or ideological options. It also facilitates mobilization because it makes clear that mobilization can be successful. Taking part, which means investing time and energy, in collective action does not make sense if one has the feeling that only a very limited number of people will take part in the effort. Within a micromobilization context, however, actors can assess the likelihood that others will participate and can even try to convince others that participation is a viable option. These networks and mobilization contexts, therefore, allow for a more reliable assessment of the chances of success of the collective action effort (Diani & McAdam, 2003). The embeddedness in preexisting networks also diminishes the propensity to act as a *free rider*, that is, to profit from the efforts of social movement organizations without sharing the burden of investing time or money in these movements.

Research has also highlighted the fact that mobilization efforts have to overcome various barriers simultaneously. First of all, it is necessary to build a general social support for the goals and ideology of the social movement. The cultural frames espoused by the social movements have to be accepted as legitimate by at least part of public opinion. Subsequently, however, social movement organizations also have to ensure specific movement support: From those who agree with the basic ideas of a movement, participants have to be recruited who are able and willing to take part in specific actions. Participants' willingness to participate depends strongly on whether the participants think these actions can really contribute to achieving the general goal of the movement (Klandermans, 1997). Klandermans himself demonstrated the validity of this approach by a study on the peace movement: Although a vast majority of the general population would agree that peace is an important goal in society, a much more limited number actually took part in activities organized by peace movement organizations.

While resource mobilization theories emphasized the rational motives of social movement participants, in the more recent literature various authors have questioned this assumption by highlighting the emotional meaning of participation (Goodwin, Jasper, & Polletta, 2001). Participation in social movements often starts from indignation or rage about a lack of social justice, even if these pioneers cannot imagine that their actions will succeed. This kind of research highlights the fact that social movements are often pioneered by people who would not even think about taking initiatives if they were motivated solely by a rational calculation of costs and benefits.

Studies focusing on the form of social movements and organizations have recently come to pay attention to the geographical scale of activism. In the 19th century, social movements tended to occur mainly within the framework of the nation-state, but this is no longer the case. As various forms of political decision making have moved toward international organizations or ad hoc gatherings of world leaders, social movements have started to focus on *transnational activism*, uniting participants and social movement organizations from various countries (della Porta, Andretta, Mosca, & Reiter, 2006). The protest surrounding a meeting of the World Trade Organization in 1999 in Seattle, Washington, is often seen as the starting point of this kind of new transnational activism. Since that time, transnational activism has been strongly present on issues such as international trade agreements, sustainable development, greenhouse gas emissions, peace, and human rights. An innovating feature of transnational activism is the emergence of a social justice movement that aims to combine various topics, such as discrimination, human rights, and environmental degradation, into one comprehensive justice framework. International organizations, however, reacted swiftly to this new form of disruption, on one hand by allowing nongovernmental organizations to have a greater say in the preparation process of international gatherings,

but on the other hand by tightening security or by even relocating summit meetings to conference sites that are difficult for protesters to access. Although at the start of the 21st century, various authors expressed high hopes about the future of transnational activism, a decade after Seattle it seems that expectations have sobered. It has become clear that this kind of activism is difficult to sustain, even as new information and communication technologies have made it easier to mobilize participants across borders. Furthermore, transnational activism itself is plagued by strong forms of inequality, in that participants from industrialized countries have more resources to participate in this kind of activism than do participants from developing countries. Nevertheless, it can be observed that the presence of nongovernmental organizations and protesters has become a fixed feature of many international conferences, and some studies have even documented the impact of this kind of social movement activity with regard to human rights, social causes, international trade agreements, and the protection of ethnic minority rights, among others. Within transnational activism, some organizations stress the fact that political decision makers should be targeted (whether at the level of international organizations or at the level of the nation-state), whereas other organizations try to circumvent political institutions by promoting lifestyle behavioral changes or by trying to apply direct pressure to international corporations (e.g., on environmental norms or child labor).

Not just on an international level, it is clear that social movements have had, and still have, strong policy effects. During the past decades, political systems have implemented profound changes with regard to civil rights, equal opportunity, peace, sustainable development, and other issues that are being championed by social movements. It has been argued, therefore, that social movements clearly matter in the political process (Giugni, McAdam, & Tilly, 1999; Meyer, 2006). Indeed, some of the major processes of democratization in the 20th century have been implemented partly as a result of the demands of social movements. From an analytical point of view, however, it is difficult to assess whether these policy changes really can be attributed to the activities of social movements. Or to phrase it in a counterfactual manner, would political systems have adopted the same reforms in the absence of social movement activity? It is almost impossible to answer this kind of question in an unequivocal manner. It is crucial, however, that social movements in any case have had a strong agenda-setting effect, bringing to the front of the political agenda topics that were considered not politically relevant. Some studies have clearly demonstrated this agenda-setting effect, such as with regard to abortion rights and children's rights. There is less agreement, however, on the exact contribution of social movements to further phases in the policy cycle. Once an issue is accepted on the political agenda, the executive, the parliaments, and the political parties tend to take control of it, which renders it

more difficult for social movement organizations to try to influence its further development.

Some studies have also highlighted the fact that “political success” in the context of social movements is not an unequivocal concept. As the issues that have been brought forward by social movement organizations are integrated into the policy agenda, movement organizations may be seen as superfluous and may become caught in a downward mobilization spiral. Although some organizations might react to this form of policy success by a process of radicalization that leads to new demands on the political system, the mainstream effect is usually demobilization and/or institutionalization. For the women’s movement, for example, it has been argued that feminist groups may no longer be as active as they were in the 1960s and 1970s, but that they are becoming increasingly intertwined with the state bureaucracy, leading to the phenomenon of *state feminism* (Stetson & Mazur, 1995).

Furthermore, however, it has to be remembered that the impact of social movements does not remain limited to changes in legislation and policy. It has also been argued that social movement activity is associated with strong cultural change, such as with regard to more attention to the environment or biodiversity, or with stronger sensitivity to injustice, such as in race or gender relations. Some feel that the women’s movement in particular succeeded in bringing about important cultural changes, leading to stronger support for equal gender roles throughout society. Survey research makes clear that issues that were pioneered by the social movements of the 1960s are now accepted by a large part of public opinion in Western countries. Some research has even pointed to a totally different form of consequences of social movement activity. For the participants themselves, participating in social movements has profound biographical consequences, leading to a pattern of continued involvement and a strengthened sense of empowerment and efficacy. Even decades after their initial involvement, social movement participants are still characterized by significantly higher levels of participation, knowledge, and efficacy (McAdam, 1988). The impact of social movements goes even further in this respect, as participants’ activities allowed for the development of distinct social identities, based, for example, on a feminist identity or a homosexual or transgender identity.

Future Directions

Since the 1970s, social movement studies have grown into a fully mature subfield straddling the boundaries of sociology and political science. It is interesting to note that social movement studies have also established links with the disciplines of social psychology, communication, international relations, and gender studies. Social movement participation is now fully accepted as an integrative part of the political action repertoire of ordinary

citizens. Within the academy, social movement studies has become an accepted subfield, with numerous courses on social movements and two well-established journals. *Mobilization. An International Journal* (established in 1996) and *Social Movement Studies* (established in 2002) continue to publish state-of-the-art research on social movements. Whereas the former journal tends to emphasize U.S. studies, the latter is more strongly rooted in the European tradition.

From a theoretical perspective, too, it can be argued that social movement studies have become a fully established and consolidated area of study. The main theoretical approaches that were developed in the 1970s and 1980s are still clearly present in the current scholarly work, and although the original insights have been qualified to a large extent, it is difficult to pinpoint any major theoretical innovations that have occurred in the past decade.

It has been quite clearly documented that social movements tend to follow a *protest cycle*, with peak periods of mobilization, followed by longer periods of demobilization. Generally, the 1960s and 1970s are considered a very intensive mobilization period, but this kind of protest behavior seems less predominant in the early years of the 21st century. To cite but one example, the war starting in Iraq in 2003 led to some mobilization by peace movement organizations, but in no way can these protests be compared to the intensive mobilization against the Vietnam War in the 1960s and 1970s. For social movement scholars, the challenge is to demonstrate that their field remains theoretically and socially relevant, even in periods when mobilization remains at relatively low levels. This challenge has led, for example, to a call for a renewed link between social movement studies and social theory. Walder (2009) has argued that social movement studies have tended to focus too narrowly on the process of mobilization, thus limiting their focus on the structural social change that is considered conducive for the rise of social movements. According to Walder, a wider focus would allow social movement scholars to relate their work to basic theoretical questions of social order and social change.

In the absence of major theoretical innovations, the focus of current research projects (which should lead to some major publications in the years ahead) is rather to apply current insights to new places and movements. Leftist and progressive movements in western Europe and the United States have been well documented by now, but if the theoretical approach to social movements really is comprehensive, it should also apply to other organizations. First of all, the scope of organizations has widened. In recent years, a number of scholars have tried to apply the insights of social movement studies to movements that thus far did not receive all that much attention in the field. Conservative or right wing associations in principle should be confronted with the same problems of mobilization, framing, and gaining access to the political decision-making

process. Although a number of studies on extreme right activism are available, however, it still remains to be investigated whether the framework that has been developed for social movements in general can also be successfully applied to this kind of social movement organizations.

Second, there has been a tendency to expand into the domain of political parties and other forms of political involvement. Political parties increasingly experience difficulties in mobilizing new members and in activating their existing members. Parties too are typically confronted with the organizational dilemma that it is extremely difficult to develop an internal organization that is coherent with their ideology. Whereas in the 1970s and 1980s, social movement participation was still seen as an innovating feature of the political action repertoire of citizens, this is no longer the case. New forms of participation have now come to the front, posing a new challenge to the theoretical study of participation behavior. Internet-mediated forms of participation, emotionally driven protest activities, transnational activism, symbolic challenges, and “spontaneous” or instantaneous forms of participation still need further investigation, and it has to be assessed whether these new forms of participation would have the same dynamics as social movement participation. Although it is clear that these new forms of protest succeed in mobilizing a large number of people, we do not know yet how successful they will be in exerting political influence, or in making sure their issues lead to a real policy output. This kind of initiative is often characterized by a rejection of traditional and hierarchical organizations, preferring more loose-knit networks or other egalitarian ways of organizing. This preference, however, could also make it more difficult to obtain real political leverage.

These new forms of participation clearly pose a challenge for mainstream social movement theory. This theory is focused rather strongly on the role of social movement organizations, and all these new developments could actually lead to a reduction of the role of organizations in the mobilization process. If protesters get mobilized by mass media or new communication technologies, the central role of organization is less evident. It remains to be investigated, however, whether challenges to the political system can be sustained in the absence of a solid organizational structure. Mass media might be very effective in mobilizing large numbers of people, yet it is also well-known that media tend to lose attention for issues after a limited time, moving on to the next issue at hand. Also, with regard to the biographical consequences of activism, we do not know whether socialization effects of involvement are just as strong in the absence of organized movements. Theda Skocpol (2003), for example, has expressed concern that social movement organizations are increasingly becoming professionalized. They no longer rely on the routine participation of a huge membership base but are increasingly being run by professionals who are paid by contributions of members. The involvement of most

members will remain limited to paying dues and occasionally reading a newsletter or taking part in a professionally organized activity. The democratic idea of grassroots involvement in policy-making processes therefore seems more and more out of reach.

Third, recent studies have also witnessed a growing geographic spread of the movements under consideration. Already in the 1980s some studies had appeared on social movements in the (then still authoritarian) countries of Central and Eastern Europe, but more recently studies have addressed social movements in Latin American, Egypt, Iran, and China. These studies have forced us to examine whether the concepts that were developed for liberal industrialized societies can be applied meaningfully to other kinds of society. They have also led to new challenges for empirical research because this kind of research entails new risks, both for researchers and for respondents in the research.

It could be argued that social movement studies are less innovating in the current development phase than they were 20 or 30 years ago. Simultaneously, however, there can be little doubt that social movement studies will remain an integral part of sociology and political science. On one hand, they allow political science to connect to the theoretical debate about major changes with regard to the social and political structure of society. Issues such as global justice, gender equality, environmental degradation, and sustainable development have received increasing attention in political decision making, and it can be argued that social movements at least have been instrumental in bringing about this change. For political science, too, social movement studies have created an opportunity to keep in touch with these social and cultural changes.

Second, however, it is now taken for granted that full political participation is a hallmark of any well-functioning democracy. This kind of participation is no longer limited to electoral participation but encompasses various forms of participation, and participation in social movement organization is clearly one of them.

Conclusion

Social movement studies can be considered an important subfield within political science. They almost inevitably depart from a view of society in which social conflicts are seen as an integral part of the social structure. Movements also allow for a further democratization of society, enabling grassroots activism, and they are instrumental in bringing new topics onto the political agenda. Social movement theory has highlighted the role of resources and cultural and political conditions for the success of social movement organizations. The most recent decade has been characterized by efforts to integrate and consolidate these theoretical approaches. It seems clear that social movement activity in the current age is lower than it was in the heyday of the 1960s and 1970s, and the main challenge for

social movement theory is to demonstrate that this subfield remains relevant for “mainstream” political science in periods of reduced social movement activity.

References and Further Readings

- Benford, R., & Snow, D. (2000). Framing processes and social movements. *Annual Review of Sociology*, 26, 611–639.
- della Porta, D., Andretta, M., Mosca, L., & Reiter, H. (2006). *Globalization from below: Transnational activists and protest networks*. Minneapolis: University of Minnesota Press.
- della Porta, D., & Diani, M. (2006). *Social movements: An introduction*. Malden, MA: Blackwell.
- della Porta, D., & Tarrow, S. (Eds.). (2005). *Transnational protest and global activism*. Lanham, MD: Rowman & Littlefield.
- Diani, M., & McAdam, D. (Eds.). (2003). *Social movements and networks*. Oxford, UK: Oxford University Press.
- Giugni, M., McAdam, D., & Tilly, C. (Eds.). (1999). *How social movements matter*. Minneapolis: University of Minnesota Press.
- Goodwin, J., Jasper, J., & Polletta, F. (Eds.). (2001). *Passionate politics: Emotions and social movements*. Chicago: University of Chicago Press.
- Johnston, H., & Klandermans, B. (Eds.). (1995). *Social movements and culture*. Minneapolis: University of Minnesota Press.
- Keck, M., & Sikkink, K. (1995). *Activists beyond borders: Advocacy networks in international politics*. Ithaca, NY: Cornell University Press.
- Klandermans, B. (1997). *The social psychology of protest*. Oxford, UK: Blackwell.
- Klandermans, B., & Staggenborg, S. (Eds.). (2002). *Methods of social movement research*. Minneapolis: University of Minnesota Press.
- Kornhauser, W. (2008). *The politics of mass society*. New Brunswick: Transaction. Introduction: I. L. Horowitz. (Original work published 1959)
- Kriesi, H., Koopmans, R., Duyvendak, J. W., & Giugni, M. (1995). *New social movements in western Europe: A comparative analysis*. Minneapolis: University of Minnesota Press.
- Laraña, E., Johnston, H., & Gusfield, J. (Eds.). (1994). *New social movements: From ideology to identity*. Philadelphia: Temple University Press.
- Le Bon, G. (1977). *The crowd: A study of the popular mind*. New York: Penguin Books. (Original French version published 1895)
- Mansbridge, J., & Morris, A. (Eds.). (2001). *Oppositional consciousness: The subjective roots of social protest*. Chicago: University of Chicago Press.
- McAdam, D. (1982). *Political process and the development of Black insurgency, 1930–1970*. Chicago: University of Chicago Press.
- McAdam, D. (1988). *Freedom summer*. New York: Oxford University Press.
- McAdam, D., McCarthy, J., & Zald, M. (Eds.). (1996). *Comparative perspectives on social movements*. Cambridge, UK: Cambridge University Press.
- McAdam, D., & Snow, D. (Eds.). (2009). *Readings on social movements: Origins, dynamics, and outcomes*. Oxford, UK: Oxford University Press.
- McCarthy, J., & Zald, M. (1977). Resource mobilization and social movements: A partial theory. *American Journal of Sociology*, 82, 1212–1241.
- Melucci, A. (1996). *Challenging codes: Collective action in the information age*. Cambridge, UK: Cambridge University Press.
- Meyer, D. (2006). *The politics of protest: Social movements in America*. New York: Oxford University Press.
- Meyer, D., Whittier, N., & Robnett, B. (Eds.). (2002). *Social movements: Identity, culture, and the state*. New York: Oxford University Press.
- Morris, A. D., & Mueller, C. M. (Eds.). (1992). *Frontiers in social movement research*. New Haven, CT: Yale University Press.
- Piven, F. F., & Cloward, R. (1977). *Poor people's movements: Why they succeed, how they fail*. New York: Pantheon Books.
- Skocpol, T. (2003). *Diminished democracy: From membership to management in American civic life*. Norman: University of Oklahoma Press.
- Stetson, D. M., & Mazur, A. (Eds.). (1995). *Comparative state feminism*. Thousand Oaks, CA: Sage.
- Tarrow, S. (1998). *Power in movement: Social movements and contentious politics*. Cambridge, UK: Cambridge University Press.
- Tilly, C. (1964). *The vendee*. Cambridge, MA: Harvard University Press.
- Tilly, C. (1978). *From mobilization to revolution*. Reading, MA: Addison Wesley.
- Tilly, C. (2008). *Contentious performances*. Cambridge, UK: Cambridge University Press.
- Tilly, C., & Wood, L. (2008). *Social movements, 1768–2008*. Boulder, CO: Paradigm.
- Walder, A. (2009). Political sociology and social movements. *Annual Review of Sociology*, 35, 393–412.

28

GENDER AND POLITICS

KAREN J. VOGEL

Hamline University

In 1976, Peter Merkl observed that the field of comparative politics had been woefully deficient with respect to the study of women (Merkl, 1976). Few comparative studies on gender existed, almost no presentations or panels appeared at professional meetings, and no academic journal specialized in the publication of research in the subfield. More than 40 years later, the study of women, and more broadly gender, in comparative politics has flourished, becoming an important area of research. A recent issue of *Perspectives on Politics* dedicated a whole section of the journal to a review of comparative politics of gender. This chapter is intended as an overview for students who wish to gain a general understanding of the evolution of this field of study.

Several aspects of the study of women in comparative politics are summarized in this chapter. The first section highlights how the comparative study of women and politics has evolved since the 1970s, noting an increase in the number of scholars in the field and the acceptance of this area of research into mainstream political science journals. The second section examines some of the key themes and theories, including women and representation and feminist comparative public policy. The growth of studies on gender regimes and the welfare state, state feminism, the formation and implementation of women-friendly policies, and the influence of women's movements on policy debates are presented. Next, some practical implications of these studies are noted, followed

by a discussion of future directions of research. Finally, the chapter concludes with a brief summary, a short list of related chapters in this handbook, and a list of references and readings for further study.

Evolution of the Study of Women and Comparative Politics and Strategies for Research

Comparative political science has been around since the time of Aristotle. In its modern configuration, the subfield has been defined as involving both a comparative method of study and a substantive area focus on understanding the societies and politics of countries and regions of the world (Hull, 1999). In its early stages, comparative politics took a more formal-legalistic approach, using historical analysis or descriptive studies of political institutions and governments primarily in Western Europe and the United States. However, with the growth of behavioral perspectives and critiques in the 1950s and 1960s, comparative studies moved toward more empirical methods and theory building (Hull, 1999). Comparative studies of political behavior, political culture, democratization, development, public policy, and the state became more common and more global, with no single theoretical paradigm dominating the field (Hull, 1999).

Even with these expansions in themes and methods, women were still remarkably absent, both as a subject of study and as practitioners in the field early on (Gruberg, 1999). Part of the reason was that gender was considered at the time to be a “marginal” or unimportant area for comparative political study. Since so few women were in key positions of political leadership around the world, the notion of women having power or political influence was perceived as irrelevant. Even the notion of gendered policy making was not considered, much less fully articulated or well developed, in comparative perspective. Furthermore, women entering political science were counseled to be careful to do “mainstream” work in order to be successful in the completion of a dissertation or accepted in the job market. Very little incentive existed for opening up new areas of research related to women and comparative public policy or political representation, or gender and politics more broadly (Hull, 1999).

By the late 1970s, this began to change as political science, and comparative politics in particular, was impacted by feminist critiques and increased numbers of women in the discipline (see also Chapter 41, titled “Feminist International Relations,” in this handbook). Changes within the American Political Science Association itself, the role of key conferences in providing the impetus for research, and the creation of new academic journals all helped the study of women and comparative politics to become accepted in the discipline.

For example, the seminal conference “Social and Political Change: The Role of Women” held at the University of California, Santa Barbara, in 1974 called attention to the need to incorporate women into political and social studies (Merkl, 1976). The formation of the Women’s Caucus for Political Science in 1969 provided mentoring and support for women entering all subfields of political science. And finally, the creation of several key academic journals, including *Signs* in 1975, *Women and Politics* in 1980 (known since 2004 as *Women, Politics, and Public Policy*), and most recently *Politics and Gender* in 2005, all fostered interest in research on women and politics. By the late 1990s, mainstream journals such as *Comparative Politics*, *Comparative Political Studies*, *PS: Political Science and Politics*, and *Perspectives in Political Science* all were publishing articles in which women were central in the analysis of various comparative topics. By the 21st century, research on women and comparative politics was much more widespread and acknowledged in political science than previously reported by studies of the discipline at the end of the 1990s (Gruberg, 1999; Hull, 1999).

Over time, multiple methodological and theoretical debates also emerged. Much like the debates in political science as a whole, the debates surrounding the comparative study of women and politics, and more recently gender and comparative politics, have focused on whether to do large cross-national, quantitative studies or to employ smaller, detailed case studies (Mazur & Perry, 1998). The

advantage of small qualitative studies is their ability to capture the complexities of cultural contexts and the interplay of race, class, and gender issues. The obvious disadvantages are the inability to generalize about cross-national patterns or to replicate particular research findings over time. On the other hand, larger, quantitative studies using data sets can offer more generalization but miss the nuances of difference that can be key in understanding the impact of identity and gender issues in particular.

Since the late 1990s, some scholars have chosen “not to choose,” that is, to try to find some combination of qualitative and quantitative methodologies such as those created by the Research Network on Gender Politics and the State (RNGS; Mazur & Perry, 1998). This network brings together more than 46 scholars to study and explore theories about gender, policy making, and state processes. Publishing a variety of books and articles over the past 10 years, these scholars have addressed some of the thorny problems of depth versus generalizability by using a combination of detailed interviews, data collection from numerous sources, numerous case studies, and more in empirical research.

Themes, Theories, and Approaches of Comparative Research on Women and Politics

Since the advent of comparative research on women and politics, several areas of inquiry have emerged, only a few of which will be discussed here. These include studies of women and political representation and feminist comparative policy, analyses of gender regimes and the welfare state, state feminism, policy formation and implementation, and the impact of women’s movements. Although women, or more broadly gender, have been commonly used as a variable rather than an analytical tool, more recent theoretical work in comparative research, overlapping with feminist international relations and women’s studies, has started to examine the possible gendered nature of policy, bureaucracies, and the state (see also Beckwith, 2010).

Women and Political Representation

The earliest comparative studies of women and politics focused first on issues of female representation in industrialized countries or the states of Western Europe, Canada, and the United States. Researchers were primarily interested in whether women were elected to key political positions and what circumstances tended to favor the access of women to political office. Formal politics, such as voting, elections, the role of political parties, and work of legislators were at the heart of these studies, with later work considering the role of women in civil society. Early classic works in these areas were *Politics and Sexual Equality: The*

Comparative Position of Women in Western Democracies (Norris, 1987), *Gender and Political Parties* (Norris & Lovenduski, 1993), and *Women and Politics Worldwide* (Nelson & Chowdhury, 1994).

In the past 15 years, studies on representation became more complex, focusing on what combination of factors contributed to a lesser or greater degree to the legislative representation of women. Political or institutional, cultural, and socioeconomic explanations also emerged as possibilities for investigation. Research also gradually moved beyond industrialized states to central and eastern Europe, Latin America, Asia, and Africa (see Kenworthy & Malami, 1999).

For example, institutional explanations looked at the nature of the electoral system itself, the construction and size of political parties and party candidate lists, the timing of women's access to the vote, and levels of democracy. Although some studies disagreed, generally proportional representation (PR) in parliaments was viewed as more favorable than district-based systems as the number of seats available to women in PR systems tended to be higher (Matland & Studlar, 1996). Political parties with open party lists and term limits were also seen as favoring the advancement of women (Htun & Jones, 2002; Krook, 2005). And overall, the quality of democracy—that is, systems with multiple, competitive political parties, high voter turnout, and regular elections—also increased the representation of women (Lindberg, 2004; see also Viterna, Fallon, & Beckfield, 2008).

A second area of interest focused on religious or cultural factors in determining levels of representation of women. Initial studies on cultural or religious factors ranged from examining the effects of Muslim or Catholic cultures in hindering women's election to political office to studying aggregated attitudes toward gender equality (Tripp & Kang, 2008). Some research attempted to use binary variables measuring whether a state ratified the United Nations (UN) Convention on the Elimination of Discrimination Against Women to examine attitudes about women's equality and the connection between these attitudes and women's representation (Tripp & Kang, 2008). More recent studies, including those by Inglehart and Norris (2003), have largely used data sets such as the World Values Survey, with information from more than 70 countries, finding that well-educated, less religious, and single respondents in postindustrial states were most likely to support equal representation for women.

Pursuing a third area of interest, scholars have explored socioeconomic factors affecting women's representation, in particular women's levels of education and participation in the labor force. These studies have yielded mixed results. Paxton (1997), for example, found that levels of education or numbers of women in the workplace are not as important as institutional factors that may limit women's access to political processes. On the other hand, Kenworthy and Malami (1999) discovered that the type of

professional occupation indeed might enhance women's participation. More recent research has returned to the question of whether development impacts women's representation. Viterna, Fallon, and Beckfield (2008) have argued that in order to have a better understanding of this relationship, taking into account different political systems and economic circumstances, studies must examine developed and developing countries separately. Levels of democracy may still be important, but economic development does matter.

In the past 6 years, researchers also explored the impact of quotas and to what degree they enhance women's representation around the world. These studies call into question earlier institutional, cultural, or socioeconomic explanatory variables (Tripp & Kang, 2008). Quotas refer to a prescribed number of seats in parliament, on political party lists, or in political leadership allocated to women within a political system. Quotas can be mandated informally through party practice or tradition, or more formally through legislative rules, constitutional amendments, and the like. Recent books on the impact and effectiveness of quotas include *Representing Women? Female Legislators in West European Parliaments* (Diez, 2005) and *Women, Quotas, and Politics* (Dahlerup, 2006). Articles on the subject have also appeared in several journals, including most recently a comprehensive study on the global impact of quotas by Tripp and Kang (2008) and work on quotas in Latin America by Jones (2009), both in *Comparative Political Studies*.

Tripp and Kang's (2008) study is particularly instructive in pointing to the impact of quotas vis-à-vis previous understanding of female representation. The nature of the electoral system, specifically in terms of proportional representation or levels of democracy, was earlier considered most important in determining the level of representation of women. However, the introduction of quotas in a variety of systems, democratic or not, has shown that although PR systems may still be important in some parts of the world, the existence of quotas, regardless of the type of regime, may be even more critical.

Quotas also seem to matter more than religious or cultural factors as many predominantly Muslim countries have adopted quotas for women, raising their representation in the political process. This phenomenon contradicts earlier research pointing to the restrictive nature of religions such as Islam or Catholicism on women's political rights. Furthermore, the existence of quotas may also be more important than the length of time women have had the vote in a particular country or the country's degree of economic development. Quotas, according to Tripp and Kang (2008), seem to have helped women overcome limitations traditionally created by economic underdevelopment, authoritarian regimes, cultural contexts, or the features of regular electoral politics, thus challenging previous notions of whether political, cultural, or economic factors are most important in understanding the evolution of women's representation.

What is more uncertain is whether quotas create a permanent avenue for representation of women and greater impact in the policy-making process. As Tripp and Kang (2008) admitted, more longitudinal studies including future data on women's representation and effectiveness in the political process are needed to learn whether quotas are indeed the key to ensuring political equality for women. Further work is also needed on understanding which women are privileged in the quota selection process. In other words, to what degree do race, class, and gender, or issues of *intersectionality* and *positionality*, matter? Additionally, more comparative work on women's representation in developing areas, especially Africa, would be helpful in understanding the possibilities for social change.

Feminist Comparative Policy

Along with women and representation, a key area of research is the study of feminist comparative policy. This research has coincided with the emergence of various new social policies, ranging from equal job opportunities, shared part-time work, welfare programs, and parental leave to policies dealing with reproductive rights, sexual harassment, and violence against women in states globally. Feminist comparative policy research is interested in understanding the similarities and differences in the formation, implementation, and effects of these policies on women, and the area is moving from studies at the state level to examinations of the impact of policies of transnational organizations such as the European Union or the UN (Gottfried & Reese, 2003). The underlying hope of researchers here is that scholars and practitioners worldwide can learn from each other what policies pragmatically solve public problems and better the lives of both men and women.

In the 1990s and early 21st century, much of this work focused on the West or women in the industrialized states. Ackelsberg (1992) found that several authors made contributions, including Gelb and Palley (1987), Sassoon (1987), and Boneparth and Stoper (1988). They explored how policies can affect men and women differently as well as to what degree women impact the policy-making process. This work raised important questions, such as: Who benefits and who does not when public problems are defined a certain way? Can women who are often marginalized or lack accessibility to decision makers affect policy? And how are inequalities reinforced, changed, or redefined by policy reform?

By the 21st century, the RNGS group, among others, began making major contributions to this research, expanding the topics and theoretical issues pursued (Mazur, 2002). Gottfried and Reese (2003) found in their comprehensive review of feminist comparative policy research at least four categories or aspects of policy study that have appeared recently in the literature. These include gender regimes, policy, and the welfare state; the emergence of state feminism and its effects; the process of women-friendly policy

formation and the challenges of policy implementation; and the role of women's movements in shaping social policy debates. Some of the themes and findings in each of these areas are discussed next.

Gender, Policy, and the Welfare State

Since the 1990s, scholars in women's studies, sociology, and political science have all been interested in trying to understand how gender is constructed in the politics of the welfare state. This work first moved toward reframing old welfare state typologies and theories into more gender-sensitive models, and then sought new ways of examining questions related to concepts of the nature of families, work, social care, and the state (Gottfried & Reese, 2003). Key to this work have been the contributions of Walby (1999), who examined the nature of gender regimes, Acker (2006), who introduced the concept of inequality regimes, and Lewis (1992), who elaborated on the concept and role of the male breadwinner model in social policy.

Gender regimes refer to systems or patterns of gender equality or inequality in gender relations found in the household, the market, civil society, and the state. Walby (1999) was interested in how gender regimes in industrialized states move from private (within the family or the personal) to public domains (economics, civil society, policy, and government). She argued that four aspects were important in understanding gender regimes and the possibility for these transitions: (1) Gender regimes were social systems or gender orders coexisting with differing degrees of inequality or equality; (2) such regimes can be differentiated in a variety of ways along market, regulatory, or socially driven dimensions; (3) a number of public domains can exist where gender regimes operate simultaneously; and (4) gender regimes can also be created as a series of gendered social activities and traditions, constructed, reconstructed, and reinforced from one generation to the next. Gender regimes do not have to be static and can change over time.

Acker (2006), a sociologist, took these ideas a step further and posited the notion of *inequality regimes* as an analytical approach to explain the creation and maintenance of inequalities in work organizations. Drawing on the concepts of intersectionality and the mutual and reinforcing reproduction of class, race, and gender relations of inequality (see O'Loughlin, Converse, & Hoeschst, 1998), she set the stage for the possibilities of detailed and complex comparisons of inequality, identity, and power relations in a variety of forms within the welfare state and beyond.

Lewis (1992) focused on another aspect of gendered policy making by exploring the importance of the *male breadwinner model*. This model contains several assumptions about the role of men in the family and the household, including the idea that men are the primary providers. Lewis, and later Orloff (2002), found that this model has often permeated assumptions of policymakers and gendered the creation of social policy across industrialized

states. Women become caught in situations in which welfare benefits or labor policies are determined by gendered assumptions that do not apply to the realities of their day-to-day lives. And even though the breadwinner model has waned or shifted with societal changes in different countries, it still carries a great deal of political power (Orloff, 2002).

Armed with these theoretical frameworks, scholars in feminist comparative policy have examined public policies related to job training, employment, education, paid work, welfare, and many other areas. They have also investigated the institutional arrangements and political contexts that have fostered these policies. While the focus has generally been on case studies of industrialized states, recent research has started to explore labor and welfare policies and institutions in central and eastern Europe, Asia, and Latin America (Pribble, 2005).

One of the interesting findings in much of this work is that social, political, and economic changes, or the move from private to public gender regimes, are not always matched by new definitions of gender roles or greater economic, social, or political equality for women (see Pascall & Lewis, 2004). As will be discussed later, only when active women's movements, creating fundamental shifts in perceptions and calling attention to equal treatment, are present do policies seem to be implemented and more likely to change conditions of inequality (Gottfried & Reese, 2003; Katzenstein & Mueller, 1987; Keiser, 1997; Montoya, 2004; Sainsbury, 1999).

State Feminism

Another important area of comparative research on women and politics is the study of state feminism. *State feminism* refers to the emergence, within governments, of women's policy machineries created to handle the advancement of women or women's policy issues. These machineries can take many different forms, such as formal ministries, offices within a particular department of the state, or special formal commissions or committees created by the parliament or legislature. According to the UN Commission on the Status of Women, women's policy machineries are any formal entity recognized and supported by the state as dealing with the promotion of equality for women. They usually are central, policy-coordinating units (see, e.g., the following UN document related to the Commission on the Status of Women: E/CN.6/1988/3; see also E/CN.6/2009/15). By 2009, two thirds of the world's states had created women's policy agencies and continued to play an important role in articulating issues related to women and girls (see, e.g., the following UN document related to the Commission on the Status of Women: E/CN.6/2009/15).

Feminist comparative research has used a variety of strategies in examining state feminism (Squires & Kantola, 2008). Approaches have included single-country case studies, multiple case comparisons selecting most similar

cases, and comparisons of case studies considered to be most different. Recent studies have moved from industrialized states to developing countries. Regardless of the number of policy cases or types of states, researchers have generally been interested in whether these machineries have an impact on women-friendly policy formation and are effective in improving the lives of women. Success might be measured in terms of whether women or advocacy groups were actually brought into critical policy-making processes and whether a redefinition of policy goals and practices occurred that addressed feminist goals of equality (Squires & Kantola, 2008).

Academic studies done by the RNGS group, as well as research initiated by the UN Division for the Advancement of Women and the UN Commission on the Status of Women, have produced a complicated picture of the impact of women's policy machineries in various regions of the world. For example, Mazur (2002) acknowledged the role of *femocrats* (a term coined by Australian feminists to describe the professional female workers in policy offices who supported gender equality agendas) in bringing feminist issues into the public discourse and the realm of formal politics. Weldon (2002) later explored the possibilities of a continuum of government responsiveness, arguing that the existence of women's agencies did not necessarily mean that they were always effective.

More recently, Outshoorn and Kantola (2007) have asked why, almost a decade after the proliferation of women's policy machineries, more progress has not been made on the achievement of equality for women. They found that often *gender mainstreaming* efforts triggered the creation of a single women's agency but at the expense of more widespread policy initiatives and funding for programs across government departments. The slowing of policy implementation also may have been caused by a need for expertise that was hard to find or by difficulty in addressing the needs of the multiple identities of women and their experiences. Furthermore, the larger economic situation and specific tensions between feminist policy goals and neoliberal market agendas collided in such a way that parties of the political right would attempt to reduce or to hold the line on costs of new social programs promoted by these agencies.

Recent work by the Division for the Advancement of Women, under the auspices of the UN Commission on the Status of Women, has also acknowledged that women's agencies must address a wide variety of issues and be supported financially and politically by governments if they are to be more effective. Recently, special attention has been given to the role these agencies may play in the fight against HIV/AIDS and the need to enhance their capacities to address education, prevention, and health issues. However, the UN itself acknowledges that without appropriate staffing, levels of financial support, and political will on the part of states, these tasks will be difficult (see, e.g., the UN document related to the Commission on the Status of Women: E/CN.6/2009/15).

Nearly all these studies agree that the success of women's policy machineries depends more on external factors than on the internal features of the bureaucracies themselves (Squires & Kantola, 2008). In other words, the general economic, social, and political environment, as well as the nature and strength of women's mobilization, matter in the sustaining of these agencies and their effectiveness. State feminism, and its ability to formulate or to implement policy, arguably is susceptible to how visible, vocal, and influential women's movements might be and whatever shifts or changes might be occurring within a country or region.

Policy Formation and Implementation

Related to the study of state feminism is an interest in policy itself and understanding the similarities and differences of policy formation and implementation cross-nationally. Along with the work of the RINGS group, several institutes and organizations have appeared in the past 30 years to extend the comparative research on women and policy. Among the most active and well-known groups are the Institute for Women's Policy Research, which works in affiliation with the George Washington University; the Center for Women Policy Studies, a nonprofit based in Washington, D.C.; and the Institute for Research on Women, located at Rutgers University. Although the work of these groups has often focused on internal policy making and implementation in the United States or Europe, interest has increased in the formation and implementation of policy for the advancement of women in developing countries as well. Many of these institutes have sponsored recent symposia, research projects, and special conferences related to the study of women and politics in Latin America, Asia, and Africa.

Scholars associated with these groups have published numerous studies using case studies, cross-national statistical analyses, or both, and have drawn a variety of conclusions. As early as 1997, Keiser found that legislation of a policy does not guarantee implementation or enforcement. Using child support policies as a case in point, Keiser noted that only where women have a strong political base and can push for accountability can policies achieve appropriate levels of enforcement (Keiser, 1997).

In their review of workplace policy research, Gottfried and Reese (2003) found multiple causal factors in effective policy formation and implementation. In particular, rates of women's unionization, types of legal systems, modernization of gender regimes, changing social and political divisions of power and authority, economic markets, and the pressures of transnational feminist networks—to name just a few—could all impact policy debates and actions to varying degrees. Gottfried and Reese also noted that further research needed to be done to clarify the role of policy recommendations and guidelines promoted by the UN and the European Union in enhancing the strength of policy implementation.

In 2004, Montoya asked theoretical questions about what makes some policies harder to implement than others and called for more complex models of research for understanding policy formation and implementation. Starting with the acknowledgment of societal values, she built a two-dimensional model focusing on various types of mechanisms needed for policy formation and the potential strategies and resources required for women's rights advocates to act and to interact with their political environments to ensure policy implementation. Focusing on the Italian experience as a case study, she concluded that depending on the policy issue, policy advocates both inside and outside government, with the potential for alliances and coordinated efforts, can affect policy outcomes.

In 2006, Hannan took up the challenge of assessing some of the broad implications and outcomes of the frameworks and recommendations for gender policy promoted by the UN Platform for Action from the Beijing Conference. She found that whereas numerous women's agencies and national policies had been created for the advancement of women and girls, the results were indeed quite mixed. Disturbing gaps continued between policy intent and policy outcomes. No matter whether it was health care, education, or the reduction of violence against women, progress was very uneven from state to state and region to region. Discrimination and public attitudes had not necessarily changed or improved at the same rate as reform efforts in legal or policy frameworks (Hannan, 2006).

Current research explores not only types of policies and their outcomes but also the importance of distilling issues related to the influence of women's movements, feminist activism, and gender issues more broadly defined. For example, Chappell (2007) has called for a shift from an agenda of comparative research on women and politics to an understanding of the role of gender norms and gendered institutions in policy making. In March 2010, Beckwith reviewed the logics of a comparative politics of gender, briefly enumerating the benefits of a comparative study of gendered analysis of political phenomena (Beckwith, 2010). This research has led to work on gender and courts, as well as opening up the possibility for studying both men's and women's positions in political structures at all government levels and in policy discourses.

Women's Movements

Since the groundbreaking work of Katzenstein and Mueller in 1987 on consciousness, political opportunity, and policy making, scholars have explored the variety of ways that feminist goals and practices connect to policy, the state, and other political structures (Katzenstein & Mueller, 1987). While Mazur (2002) cautioned that the role of women's movements is not always clear-cut, most feminist comparative studies have agreed that some level of women's mobilization, through either autonomous

movements or larger formal organizations such as unions and political parties, has been key in the formation and effectiveness of women-friendly policies in Europe and Latin America (DiMarco, 2008; Kahn & Meehan, 1992).

Recent literature has continued to pose complex questions surrounding the evolution and role of women's movements in civil society in different cultural contexts. Some of this scholarship has also explored the strength or viability of cooperation among local women's nongovernmental organizations and global networks in influencing gender policy. For instance, Orr (2008) found that in some east European states, ethnic and regional divisions limited the ability of women's organizations to work together on issues related to equality. Comparative scholarship on women and politics in the Middle East pointed to positive relationships between transnational women's networks and global rights agendas promoted by the UN and women's movements, as well as the important role women have had in shaping civil society (Moghadam & Sadiqi, 2006).

Research shows that women's activism does not necessarily reproduce Western gender frameworks but involves pragmatic and context-specific strategies that are still responsive to global trends promoting the advancement of women (Moghadam & Sadiqi, 2006). Although more studies need to be done regarding trends in Africa, early work suggests that here, too, cultural context matters and that strategies to improve women's lives may not depend only on influencing formal structures of political systems but rather on the creation of local and communal women's organizations and networks (Goetz & Hassim, 2003; Lindberg, 2004).

Practical Implications

Practical implications for policy development and research can be drawn from the work on women and comparative politics over the past several decades. First and foremost, context matters; policies formed in the United States on pay equity and parental leave, for example, may not work or even be appropriate for countries in Asia or the Middle East, because of differing histories, cultures, economic systems, gender regimes, and political structures. Even though scholars and practitioners have much to learn from each other, importing gender policies from one country to the next should be considered with great caution and understanding that circumstances and conditions can differ from state to state.

At the same time, that wisdom does not mean that general guidelines on the equal treatment of women should not be discussed or that general theorizing and empirical research should be abandoned. The creation of universal policies for all may not be possible, but international bodies such as the UN, the European Union, the African Union, and other transnational networks have roles to play in drawing attention to the condition of women worldwide and providing recommendations, support, and

possible strategies for improvement where appropriate. Furthermore, the comparative study of similarities and differences in policy making can provide insights for policymakers within states, as well as point to broad patterns and themes cross-nationally. Placing gender at the center of these studies can also improve understandings of politics and political behavior generally.

The use of quotas as a practical solution for increasing women's representation in political systems is also important to note. Although further research should be done on the longevity and real influences on political processes overall, comparative research so far seems to suggest that this type of institutional reform may work at least partially to break old patriarchal norms and provide more representation to women in some contexts.

Future Directions

The study of women and comparative politics continues to be one of the dynamic areas within political science, with much research yet to be done. Only a few future directions can be suggested here. First, it is clear that more work is needed on the representation, political influences, and policy processes of women in developing countries. A paucity of research exists on women and Africa in particular. So far, most studies have focused on South Africa or Uganda, and few have gone beyond discussing issues of representation and policy formation (see Goetz & Hassim, 2003; Lindberg, 2004). Additional research on a wider variety of countries and policies would improve the understanding of diverse historical, cultural, social, economic, and political factors in the region, as well as gender activism and policy reform from state to state.

As already discussed, continued research is needed on quotas as a practical solution for the advancement of women in politics. What is the long-term impact of quotas? Does election through quotas affect the perceptions of women by their peers? Are they viewed as equally qualified, or does the quota selection process tend to color their acceptance by fellow legislators, inhibiting their performance as policymakers? What happens after quota systems are lifted? And does the existence of quotas in party lists, parliaments, or other political bodies have any spillover effect in the direction of other policies related to equal treatment either for women or for other marginalized groups? And do race and class among women also matter in who is elected through quotas? These questions might be very interesting to explore in a variety of cultural contexts.

Additional work on the process of policy transfer through international organizations like the African Union, the European Union, or the UN would also be useful. As yet, the academic literature has only started to explore to what extent states take these recommendations seriously and whether such policy recommendations are effective in improving women's lives. With respect to policy outcomes

as well as issues of representation and quotas, more longitudinal research is needed to understand the longer term impact of strategies, programs, and agendas for the advancement of women.

Shifts to focusing on gender more broadly in comparative studies also open up the possibilities for more work on lesbian, gay, bisexual, and transgender issues and public policy. To date, some comparative work has been done on politics and policies related to homosexuality and gay rights in Canada and the United States (Smith, 2008). However, considerably more research is needed to understand gender dynamics, political mobilization, and possible new directions for policy creation, reform, and implementation in this area in other countries around the world.

Conclusion

This chapter has traced the evolution of the study of women and comparative politics from the later 1970s to the present. Using methodologies common to political science, comparative research on women and politics has moved from topics focused on representation of women in various governments around the world to the study of feminist comparative policy. Over time, theoretical frames have incorporated intersectionality, gender regimes, and interests in gender as an analytical category, especially with respect to the state and policy-making practices.

Practical implications for research include the recognition that historical, cultural, social, and political contexts matter in understanding policy formation and implementation. Case studies are vital in gaining these insights, and yet cross-national studies are still important in efforts to understand broad patterns and relationships of representation. Future directions for research include the need for more comparative work on women and politics in developing countries, especially Africa, and the topic of comparative public policy and sexual diversity. It is hoped that this chapter prompts students to consider doing research of their own in this dynamic area of political science.

References and Further Readings

- Ackelsberg, M. (1992). Review: Feminist analyses of public policy. *Comparative Politics*, 24(4), 477–493.
- Acker, J. (2006). Inequality regimes: Gender, class, and race in organizations. *Gender & Society*, 20(4), 441–464.
- Beckwith, K. (2010). Introduction: Comparative politics and the logics of a comparative politics of gender. *Perspectives on Politics*, 8(1), 159–168.
- Boneparth, E., & Stoper, E. (Eds.). (1988). *Women, power, and policy: Toward the year 2000*. New York: Pergamon.
- Chappell, L. (2007). *Comparative gender and institutions: Directions for research*. Paper presented at the American Political Science Association Meeting.
- Dahlerup, D. (Ed.). (2006). *Women, quotas, and politics*. London: Routledge.
- Diez, M. M. (2005). *Representing women? Female legislators in West European parliaments*. Colchester, UK: ECPR Press.
- DiMarco, G. (2008). Social justice and gender rights (Report). *International Social Science Journal*, 59(191), 43–56.
- Elman, A. (2003). Engendering state theory: Feminists engage the state. *Review of Policy Research*, 20(3), 549–558.
- Galligan, Y. (2007). Gender and political representation: Current empirical perspectives. *International Political Science Review*, 28(5), 557–570.
- Gelb, J., & Palley, M. (Eds.). (1987). *Women and public policies*. Princeton, NJ: Princeton University Press.
- Goetz, A., & Hassim, S. (Eds.). (2003). *No short cuts to power: African women in politics and policy making*. London: Zen Books & David Philip.
- Gottfried, H., & Reese, L. (2003). Gender, policy, politics, and work: Feminist comparative and transnational research. *Review of Policy Research*, 20(1), 3–18.
- Gruberg, M. (1999). Participation by women in the 1999 APSA annual meeting. *PS: Political Science and Politics*, 32(4), 761–762.
- Hannan, C. (2006). Empowering women: Ten years after the Beijing conference (a look back). *Georgetown Journal of International Affairs*, 7(2), 173–183.
- Htun, M., & Jones, M. (2002). Engendering the right to participation in decision making: Electoral quotas and women's leadership in Latin America. In N. Craske & M. Molyneux (Eds.), *Gender and the politics of rights and democracy in Latin America* (pp. 32–56). New York: Palgrave.
- Hull, A. (1999). Comparative political science: An inventory and assessment since the 1980s. *PS: Political Science and Politics*, 32(1), 117–225.
- Inglehart, R., & Norris, P. (2003). *Rising tide: Gender equality and cultural change around the world*. New York: Cambridge University Press.
- Jones, M. P. (2009). Latin American gender quotas, electoral laws, and the election of women. *Comparative Political Studies*, 42(1), 56–81.
- Kahn, P., & Meehan, E. (Eds.). (1992). *Equal value/comparable worth in the UK and the USA*. New York: St. Martin's Press.
- Katzenstein, M., & Mueller, C. (Eds.). (1987). *The women's movements of the U.S. and Western Europe: Consciousness, political opportunity, and public policy*. Philadelphia: Temple University Press.
- Keiser, L. (1997). The influence of women's political power on bureaucratic output: The case of child support enforcement. *British Journal of Political Science*, 27(1), 136–148.
- Kenworthy, L., & Malami, M. (1999). Gender inequality in political representation: A world wide comparative analysis. *Social Forces*, 78(1), 235–269.
- Krook, M. (2005). *Comparing methods for studying women in politics: Statistical, case study, and qualitative comparative techniques*. Paper presented at the American Political Science Association Meeting, Washington, DC.

- Lewis, J. (1992). Gender and development of welfare regimes. *Journal of European Social Policy*, 3, 159-173.
- Lindberg, S. (2004). Women's empowerment and democratization: The effects of electoral systems, participation, and experience in Africa. *Studies in Comparative International Development*, 39(1), 28-53.
- Lovenduski, J. (Ed.). (2005). *State feminism and political representation*. Cambridge, UK: Cambridge University Press.
- Matland, R. E., & Studlar, D. T. (1996). The contagion of women candidates in single member districts and proportional representation systems: Canada and Norway. *Journal of Politics*, 58(3), 707-733.
- Mazur, A. (2002). *Theorizing feminist policy*. Oxford, UK: Oxford University Press.
- Mazur, A., & Perry, J. (1998). Choosing not to choose in comparative policy research design: The case of the Research Network on Gender, Politics, and the State. *Policy Studies Journal*, 26, 384-400.
- Merkel, P. (1976). The study of women in comparative politics: Reflections on a conference. *Signs*, 1(3), 749-756.
- Moghadam, V., & Sadiqi, F. (2006). Women's activism and the public sphere: Introduction and overview. *JMEWS: Journal of Middle East Women's Studies*, 2(2), 1-7.
- Montoya, C. (2004). *From rhetoric to reality? The implementation of women's rights in Italy and Western Europe*. Paper presented at the Southern Political Science Association Meeting, New Orleans, LA.
- Nelson, B., & Chowdhury, N. (1994). *Women and politics world wide*. New Haven, CT: Yale University Press.
- Norris, P. (1987). *Politics and sexual equality: The comparative position of women in Western democracies*. Boulder, CO: Rienner.
- Norris, P., & Lovenduski, J. (1993). *Gender and party politics*. London: Sage.
- O'Loughlin, P., Converse, N., & Hoeschst, H. (1998). Listening to women's voices: Toward a comparative politics of women's differences. In M. Segal & V. Demos (Eds.), *Advances in gender research* (pp. 103-137). London: JAI Press.
- Orloff, A. (2002). *Gender equality, women's employment: Cross national patterns of policy and politics*. Paper prepared for the Workshop on Welfare, Work and Family: Southern Europe in Comparative Perspective. Florence, Italy: European University Institute.
- Orr, S. (2008). Identity and civil society in Latvia, Poland, and Ukraine: Women's NGOs. *East European Politics and Societies*, 22, 856-880.
- Outshoorn, J., & Kantola, J. (Eds.). (2007). *Changing state feminisms*. New York: Palgrave Macmillan.
- Pascall, G., & Lewis, J. (2004). Emerging gender regimes and policies for gender equality in a wider Europe. *Journal of Social Policy*, 33(3), 373-395.
- Paxton, P. (1997). Women in national legislatures: A cross national analysis. *Social Science Research*, 26(4), 441-464.
- Pribble, J. (2005). Promises of empowerment: Women in Asia and Latin America (Review). *Latin American Politics & Society*, 47(4), 189-193.
- Sainsbury, D. (1999). *Gender regimes and welfare states*. Oxford, UK: Oxford University Press.
- Sassoon, A. (Ed.). (1987). *Women and the state: The shifting boundaries of public and private*. London: Hutchinson.
- Smith, M. (2008). *Political institutions and lesbian and gay rights in the United States and Canada*. New York: Routledge.
- Squires, I., & Kantola, J. E. (2008). *From state feminism to market feminism?* Paper presented at the International Studies Association 49th Annual Conference, San Francisco.
- Stevens, A. (2008). Book review: Changing state feminism. *Western European Politics*, 31, 637-638.
- Tripp, A., & Kang, A. (2008). The global impact of quotas: On the fast track to increased female legislative representation. *Comparative Political Studies*, 41, 338-361.
- Viterna, J., Fallon, K., & Beckfield, J. (2008). How development matters: A research note on the relationship between development, democracy, and women's political representation. *International Journal of Comparative Sociology*, 49, 455-477.
- Walby, S. (1999). The new regulatory state: The social powers of the European Union. *British Journal of Sociology*, 50(1), 118-140.
- Weldon, S. (2002). *Protest, policy, and the problem of violence against women: A cross national comparison*. Pittsburgh, PA: Pittsburgh University Press.

29

COMPARATIVE ENVIRONMENTAL POLITICS AND CONFLICT

AMANDA M. ROSEN

Webster University

Since the last years of the cold war, the study of security and conflict has moved forward from traditional interstate wars to consider issues of *human security*. Ullman (1983) was one of the first to call for the field of security studies to include nontraditional forms of conflict, and other scholars (Kaplan, 1994; Matthews, 1989) soon followed, noting in particular the role that environmental issues could play in creating and exacerbating conflict between states.

Connections between environmental scarcity and conflict date to the original work of Thomas Malthus (1798), who predicted that the difference in the rate of growth in population and food supply would eventually lead to Earth's population overtaking the food supply. Choucri and North (1975) argue that rising population creates an increased demand on limited resources, which in turn causes "lateral pressure" that leads countries to seek resources outside their borders via conquest. However, the field of environment and conflict really took hold in 1991 with the publication of Thomas Homer-Dixon's article titled "On the Threshold: Environmental Changes as Causes of Acute Conflict" in the journal *International Security*. This article laid the groundwork for new thinking about how environmental scarcity can lead to social changes that in turn create breeding grounds for intra- and interstate conflict.

This chapter examines the foundational ideas of Homer-Dixon and his colleagues at the University of Toronto, dubbed the *Toronto Group*; the additions and critiques offered by other scholars in the field; the empirical evidence for and against the Toronto Group's hypotheses; the policy implications of the argument; and the future directions of research in environment and conflict.

Theory

There are numerous environmental challenges that can create incentives for conflict. Issues such as climate change, deforestation, depletion of fish stocks, water pollution, ozone layer depletion, and degradation of agricultural land present clear challenges to policymakers interested in avoiding conflict. For centuries, human settlement patterns have followed the path of natural resources and their ability to provide food, water, and shelter. Changes to these environments can therefore disrupt established social patterns and lead to conflicts between groups and states. Essentially, environmental scarcity puts pressure on existing social processes, resulting in decreased agricultural production, economic decline, population displacement, and disruption of normal patterns of social relations. This potentially

leads to conflict between states and societies (Homer-Dixon, 1991).

Environmental Scarcity

Homer-Dixon (1998) defines *environmental scarcity* as “scarcity of renewable resources such as cropland, forests, river water, and fish stocks” (p. 8). There are three types of environmental stress or scarcity. *Supply* scarcity results when an existing resource is depleted, such as when deforestation reduces the availability of lumber for fuel and shelter. Demand for the resource remains the same, but there is less of it to go around. The second type is *demand* scarcity. This occurs when the resource stock remains static, but the population relying on the resource increases, through either higher levels of growth or migration. Supply of the resource remains the same, but it must be split into smaller pieces to accommodate everyone. The third type of scarcity results when the *distribution* of the resource is not equal. Supply or demand may remain unchanged, but groups may have unequal access to the resource, creating a system of haves and have-nots.

Environmental scarcity can result from several different factors. Changes in the environment can lead to degradation of an existing resource. For example, global warming could lead to rising sea levels, creating soil erosion and flooding that reduces the supply of resources in an area. Alternatively, high levels of population growth, as commonly found in developing countries, can increase demand pressures on local agricultural yields or fish stocks. Finally, government policies can create unequal access to resources, with certain groups being favored over others. The Israeli government, for example, routinely allowed Israeli settlers greater access than Arab settlers to scarce water resources in the West Bank settlements.

Social Pressures

Environmental scarcity can negatively impact social relations and economic development, which can increase tensions between states and groups and create conditions ripe for conflict. Scholars note four *social pressures* in particular that can lead to conflict: reduced agricultural production, economic decline, population displacement, and disruption of normal patterns of social relations (Homer-Dixon, 1991).

Reduced Agricultural Production

Reduced agricultural production could result from numerous environmental changes, such as deforestation, increased runoff, soil erosion, and flooding from climate change. Agricultural land located in low-lying regions such as Bangladesh is particularly vulnerable to the more intense storms and flooding expected of climate change.

Indeed, changes in precipitation and insect migration patterns could have devastating impacts on agricultural yields. Likewise, deforestation practices accelerate erosion, leaving fertile lands vulnerable to flooding during seasons of high precipitation. It is unclear whether technological developments will be able to ensure continued increases in agricultural yields when resources are faced with these kinds of environmental changes. In developing countries in particular, agriculture is a large share of the economy and the key to economic growth, and the impact of reduced production would be profoundly felt (Homer-Dixon, 1998).

Economic Decline

A second social effect of environmental scarcity is centered on economic development. Economic decline can result from numerous environmental changes, such as environmental disasters from climate change, deforestation practices hurting river trade and decreasing fuel wood, the loss of fish stocks, and increased pollution's increasing the prevalence of human disease. Overall economic development may be constrained by these changes, or even more likely, such changes could result in an increase in the wealth gap between elites and non-elites (Homer-Dixon, 1998).

Short-term efforts at increasing economic production can have long-term implications that lead to overall decline. Logging, for example, can be a productive industry in the short term, but deforestation can lower productivity over time. Likewise, increasing energy consumption is a mark of development, but the long-term implications of global warming can be very costly. Indeed, in 2006 the *Stern Review on the Economics of Climate Change* noted that halting global warming would cost around 1% of global gross domestic product—but failing to do so could cost up to 20% of gross domestic product (Stern, 2006).

Migration

A third potential impact of environmental scarcity is an increase in the number of environmental refugees (Jacobson, 1988; Westing, 1992) migrating because of environmental pressures in their home region (push) or economic opportunity elsewhere (pull). Homer-Dixon (1998) suggests that a key factor influencing migration is the gap between satisfaction in the home region and expected satisfaction in the new region. For example, migration patterns in China from rural to urban regions can partially be attributed to scarce resources in the interior leading to lower rates of economic growth. This may be due to either environmental change or, more likely, high rates of population growth, which create demand scarcities on resources. Citizens leave, not only because of the lack of resources in the home region but in order to seek opportunities in areas of higher growth.

Disruption of Social Relations

The fourth impact of environmental scarcity is a disruption of normal patterns of social behavior. "Scarcity sharpens distinctions between winners and losers—between groups that profit from scarcity and those that are hurt" (Homer-Dixon, 1998, p. 96). Groups that may have peacefully coexisted in a time of plenty become encouraged to focus on group survival, creating tense interactions with other resource-dependent groups also focused on survival. Existing divisions between ethnic, religious, and linguistic groups become reinforced, and these groups may decide to confront authority figures with violence rather than with peaceful negotiation. Social institutions such as the state itself may become fractured as a result (Homer-Dixon, 1998).

Conflict Types

These social effects may, in turn, cause several specific types of acute conflict, including scarcity disputes between countries, clashes between ethnic groups, and civil strife and insurgency, each with potentially serious repercussions for the security interests of the developed world. Three types of conflict might result from the social pressures caused by environmental scarcity.

*Environmental Scarcity**Leads to Simple Scarcity Conflicts*

Simple scarcity conflicts arise when structural conditions lead states to use violence to increase their share of the resource pie (Choucri & North, 1975). Structural theories suggest that as the power of challenger groups rises, challenger groups may choose to challenge the existing authorities with violence rather than with peaceful measures, in anticipation of moving from challenger to authority (Homer-Dixon, 1998). Simple scarcity conflicts, therefore, can be considered *resource wars*. Although conflicts over coal, oil, or other strategic resources would fit into this category, the clear focus is on wars over renewable resources. There are a number of examples of the former, such as World War I, the Falklands War, and the Persian Gulf War, but the latter is much more rare, leading the Toronto Group to downplay the role of environmental scarcity in causing this type of conflict.

*Environmental Scarcity**Leads to Group Identity Conflicts*

A second type of conflict occurs when population growth and migration create problems between groups. *Us versus them* conflicts develop as groups move from areas in decline to areas of new opportunities and interact with the inhabitants, who now must split their fixed resources among more people. These demand-type scarcities exacerbate tensions

between groups that may have peacefully coexisted in the past and interact with other social and political factors to create possibilities for violent conflict. The conflicts in Darfur and Rwanda could fall into this category.

*Environmental Scarcity**Leads to Insurgencies Against the State*

Environmental scarcity can also increase the likelihood of violent challenges to the state. As one group experiences deprivation, it will measure its access to resources against the access of other groups, particularly those holding political power. As the relative deprivation of a group not in authority increases, a sense of economic justice may lead to an increase in its grievances against the group with authority. It may look for opportunities to challenge the group in power, and circumstances may align to lead to a change in the relative power between these groups. If the disenfranchised group perceives that power relations have changed, then it may choose to address its grievances against current authorities through violence. For example, in the early 1990s in Chiapas, Mexico, economic reforms weakened the existing regime and its ability to maintain power through coercive measures. Other social institutions—notably the Catholic Church—stepped into the gap, creating social support for challenger groups and helping to create an "insurgent consciousness" (Homer-Dixon, 1998).

Critiques

The work of the Toronto Group is not without its critics. Criticism of these arguments focuses on three main areas: the definition of environmental conflict and security, a chicken-and-egg debate over whether environmental scarcity leads to or results from conflict, and the empirical evidence supporting the above hypotheses.

Defining Environmental Security

The Toronto Group's argument advocates a comprehensive view of security issues that includes environmental causes of conflict. Several scholars disagree with this approach. Frederick (1999) argues that security should be divided into two categories: traditional security issues and nontraditional problems. Deudney (1999) takes this claim further, noting that environmental security issues have causes that are fundamentally different from the causes of traditional security concerns. Traditional threats, for example, result from external aggression and provoke an us-versus-them mentality; environmental conflict, however, is a long-term threat that requires states to cooperate in order to solve it. Scholars continue to be divided over whether to treat environmental causes of conflict as a neglected element of traditional security concerns or as a completely different issue (Matthew, 1999).

The Direction of Causality

Some scholars note that environmental scarcity may actually be the result of conflict, rather than its cause (Gleditsch, 1998). According to this line of reasoning, conflicts engineered for political or economic reasons can have crucial environmental impacts, disrupting social order and economic life. These critics note that it is very difficult to tease out the individual role played by environmental factors in causing war and that current empirical research has failed to show that environmental scarcity is the cause rather than the result of conflict.

Empirical Evidence

Critics also cite the lack of conclusive empirical evidence for environmental security theories as a reason to doubt their validity. Much of the early work in the field focused on anecdotal examples rather than rigorous hypothesis testing. Later efforts by the Toronto Group attempted to address this grievance through multiple case studies, which suggest support for the theory. As noted below, the empirical evidence is decidedly mixed, with some studies (Hendrix & Glaser, 2007; Urdal, 2005) supporting claims that environmental factors increase the risk of armed conflict and others (Binningsbo, de Soysa, & Gleditsch, 2007; de Soysa, 2002) showing that the same factors actually reduce conflict risk. Indeed, many scholars remain concerned about the veracity of the Toronto Group's claims, as methodological issues persist (Buhaug, Gleditsch, & Theisen, 2008; Gleditsch, 1998).

Applications and Empirical Evidence

Much of the debate over the role of the environment in causing conflict has focused on the empirical evidence supporting the claims of the Toronto Group. Numerous case studies provide evidence of links between environmental scarcity and conflict, but these studies are often plagued by methodological concerns (Gleditsch, 1998). Quantitative approaches are more recent and have had decidedly mixed results. Most recent studies have focused largely on the issue of population growth and density and their causal role in conflict, but the other types of conflict have also received their share of attention.

Environmental Scarcity and Conflicts of Simple Scarcity

Salehyan (2008) notes that "while environmental degradation is certainly not a necessary condition for armed conflict, neither is it a sufficient one" (p. 317). Indeed, there is little evidence supporting the role of environmental scarcity in creating simple scarcity conflicts despite the many cited examples of "resource wars." The

distinction is that while wars are frequently fought over nonrenewable resources, wars fought over the renewable resources that define environmental scarcity are much rarer. Scarce nonrenewables such as oil and coal are frequently fought over, probably because they are more easily transformed into state power than renewables are.

The only scarce renewable resource that appears to strongly contribute to conflicts is river water. When a dependent downstream country relies on an upstream country to keep the flow moving along, conflict can arise should the upstream country decide to limit supply in any way. Water scarcity issues play a particularly important role in Israel and peace efforts in the Middle East. The differences in socioeconomic conditions between Israelis and Palestinians are only exacerbated by unequal access to scarce water resources (Lowi, 1999). Water has also been used as a strategic tool. For example, to compel Syria to take action against Kurd separatists, Turkey threatened to restrict the flow of the Euphrates River. But some scholars note that although water can play a role in conflicts, it is difficult to pinpoint cases in which inequalities of water distribution were a key cause of conflict (Loneragan, 2001).

Environmental Scarcity and Conflicts of Group Identity and Insurgency

Most of the empirical studies of the connection between environmental scarcity and conflict focus on group-identity conflicts, particularly those that create social instability and lead to insurgencies. Anecdotal evidence is particularly strong in this area—proponents cite, for example, the high levels of migration from Bangladesh to India as a causal factor in conflicts in border Indian states. Likewise, political instability in China could be attributed to population pressures on limited arable land (Goldstone, 2001).

Hauge and Ellingsen (1998) conducted a study showing that land degradation, freshwater scarcity, population density, and deforestation have direct positive effects on the incidence of civil war. They isolated the role of environmental causes from other sociopolitical causes of conflict and found evidence in favor of environmental scarcity's role in causing conflict. Attempts to replicate this study have been unsuccessful, however (Theisen, 2008). Tir and Diehl (2001) found a statistically significant connection between population growth and involvement in militarized interstate disputes. A study by Miguel, Satyanath, and Sergenti (2004) found that the effect of unexpectedly low precipitation levels on national economic growth in countries of sub-Saharan Africa increased their risk of civil war. Other scholars have made similar findings, demonstrating a connection between increased water scarcity or changes in precipitation and conflict (Hendrix & Glaser, 2007; Raleigh & Urdal, 2007). The connection between population density

and the likelihood of conflict has received particular attention. Scholars note that pure population growth does not capture the demand-side issue of scarcity unless density is also accounted for (de Soysa, 2002; Raleigh & Urdal, 2007; Urdal, 2005). However, some scholars have found that population density does not play a role (Tir & Diehl, 2001) and that resource wealth, rather than scarcity, is associated with an increased likelihood of conflict (Binningsbo et al., 2007).

Process tracing of numerous case studies also provides evidence for the role of environmental scarcity in provoking or exacerbating group conflict. Water issues in Gaza provoked tensions between Arab and Israeli settlers, and social turmoil in South Africa in the late 1980s can be partially attributed to unequal access to resources by Blacks and Whites (Kelly & Homer-Dixon, 1998; Percival & Homer-Dixon, 2001). In the Philippines, degradation of cropland and forests has exacerbated economic problems, and insurgency movements have been empowered by the poverty of agricultural workers operating in areas of weak central government. However, evidence in at least one case—the Rwandan genocide—shows that although environmental degradation and high population density were present, these factors did not play a major role in escalating conflict between Hutus and Tutsis (Percival & Homer-Dixon, 1998).

The empirical record for a generic connection between environmental scarcity and armed conflict is therefore decidedly mixed. As many critics have pointed out, existing efforts have failed to produce robust findings confirming that environmental factors play a role separate from other sociopolitical causes of conflict (Buhaug et al., 2008). However, the case studies and quantitative efforts do suggest that the hypotheses of the Toronto Group cannot be ignored and merit further study.

Policy Implications

The empirical evidence does suggest that environmental scarcity can increase the likelihood of certain kinds of conflict. Policymakers eager to avoid these types of conflict can take several pieces of advice from scholars.

Homer-Dixon (1991) noted that there are two paths to avoiding conflicts posed by environmental scarcity. One possibility is to become a “hard” authoritarian regime, cracking down on any social unrest that results from resource scarcity or population pressures. A second option is to promote technical and social ingenuity (Homer-Dixon, 2002): *Technical ingenuity* can lead to more efficient uses of scarce resources, diminishing the need for conflict over their use and ownership, and *social ingenuity* can create an economy and a society that do not rely on the consumption of scarce resources to survive.

If all else fails, standard military reactions to conflict are available for outbreaks of violence caused by

environmental scarcity. Butts (1999) notes that the U.S. military has extensive resources and skills that can easily be brought to bear on any conflicts that arise this way.

Environmental scarcity does pose a profound challenge to policymakers. Many policies that can lead to conflict via environmental changes are grounded in principles of economic development. For developing countries, for example, deforestation practices are designed not to destroy the environment or degrade agricultural land, but to provide fuel for the locals and lumber for trade. Indeed, practices viewed by wealthy, developed countries as degrading the environment are viewed by poorer, developing states as practical necessities for strong economic growth. It can be very difficult for policymakers to focus on the potential long-term implications of policies designed to produce short-term economic gains. Policymakers therefore must pay particular attention to the concerns of developing countries if they hope to mitigate the effects of environmental change on future conflicts.

Further challenges are posed by the nature of environmental problems. Political borders mean little to rivers, oceans, or the atmosphere. One state’s pollution can quickly become the problem of its neighbors and indeed the entire globe. For example, the greenhouse gases causing global warming are emitted by many states, but more than 40% of the total comes from the United States and China. These gases—and their effects on the warming of the globe and climate change—do not affect only the emitters, however. Their impact will be felt by every country on the planet, including those states that emit zero greenhouse gases. Collective action is required to address these causes of conflict. Therefore, as they craft international regimes promoting cooperation, policymakers should pay particular attention to designing regimes that can manage environmental causes of conflict.

Future Directions

Empirical testing of the Toronto Group’s hypotheses is still in the preliminary stages, and additional analysis is needed. In particular, scholars are beginning to pay close attention to isolating the environmental causes of conflict from more traditional political and economic factors (Hauge & Ellingsen, 2001). More research in this area in particular is needed to settle long-standing debates on whether environmental factors by themselves are likely to cause security problems and indeed whether environmental conflict should be treated similarly to more traditional sources of conflict.

As noted in the previous section, international collective action is frequently the key to addressing environmental issues. In recent years, scholarship has focused on the issue of the design of international regimes, paying close attention to how these regimes can be designed to

promote cooperation and compliance (Koremenos, Lipson, & Snidal, 2001). Scholarship on environment and conflict could benefit from a closer look at this related discipline and draw new conclusions on how to minimize the environmental sources of inter- and intrastate conflict.

Finally, scholars are starting to question the Toronto Group's assertion that climate change will be a minimal driver of conflict in the near future. Scientific assessment has accelerated the timetable by which the major effects of climate change will be felt, and the actions that policymakers take today can have a direct impact on the likelihood of conflict in the future. Many scholars have taken up the task of analyzing policy options that could mitigate the impact of climate change generally (Rabe, 2004), but few have focused their attention particularly on the security implications of climate change and how they might be avoided or alleviated. The recent explosion in work on climate change in general bodes well that this crucial issue will soon receive some much-deserved attention.

Conclusion

This chapter introduced the role of environmental change and scarcity in causing conflict. The Toronto Group's claim that environmental scarcity creates social changes that create a breeding ground for conflict led to a series of debates on where environmental issues fall along the traditional spectrum of security studies. Empirical evidence suggests that there is some support for the Toronto Group's contentions; however, further analysis is needed, as is a better conception of the term *environmental conflict*. Recent developments in the study of the design of international institutions, however, provide hope that policymakers may be able to minimize the impact of environmental change as an instigator of conflict.

References and Further Readings

- Binningsbo, H. M., de Soysa, I., & Gleditsch, N. P. (2007). Green giant or straw man? Environmental pressure and civil conflict. *Population and Environment*, 28(6), 337–353.
- Buhaug, H., Gleditsch, N. P., & Theisen, O. M. (2008). *Climate change, the environment, and armed conflict*. Paper presented at the annual meeting of the American Political Science Association, Boston.
- Butts, K. H. (1999). The case for DOD involvement in environmental security. In D. H. Deudney & R. A. Matthew (Eds.), *Contested grounds: Security and conflict in the new environmental politics* (pp. 109–126). Albany: SUNY Press.
- Choucri, N., & North, R. (1975). *Nations in conflict*. San Francisco: Freeman.
- de Soysa, I. (2002). Paradise is a bazaar? Greed, creed, and governance in civil war, 1989–1999. *Journal of Peace Research*, 39(4), 395–416.
- Deudney, D. H. (1990). The case against linking environmental degradation and national security. *Millennium*, 19(3), 461–476.
- Deudney, D. H. (1999). Environmental security: A critique. In D. H. Deudney & R. A. Matthew (Eds.), *Contested grounds: Security and conflict in the new environmental politics* (pp. 187–219). Albany: SUNY Press.
- Frederick, M. (1999). A realist's conceptual definition of environmental security. In D. H. Deudney & R. A. Matthew (Eds.), *Contested grounds: Security and conflict in the new environmental politics* (pp. 91–108). Albany: SUNY Press.
- Gleditsch, N. P. (1998). Armed conflict and the environment: A critique of the literature. *Journal of Peace Research*, 35(3), 381–400.
- Goldstone, J. A. (2001). Demography, environment, and security. In P. F. Diehl & N. P. Gleditsch (Eds.), *Environmental conflict* (pp. 84–108). Boulder, CO: Westview Press.
- Graeger, N. (1996). Environmental Security? *Journal of Peace Research*, 33(10), 109–116.
- Hauge, W., & Ellingsen, T. (1998). Beyond environmental scarcity: Casual pathways to conflict. *Journal of Peace Research*, 35(3), 299–317.
- Hauge, W., & Ellingsen, T. (2001). Causal pathways to conflict. In P. F. Diehl & N. P. Gleditsch (Eds.), *Environmental conflict* (pp. 36–57). Boulder, CO: Westview Press.
- Hendrix, C. S., & Glaser, S. M. (2007). Trends and triggers: Climate change and civil conflict in sub Saharan Africa. *Political Geography*, 26(6), 695–715.
- Homer Dixon, T. F. (1991). On the threshold: Environmental changes as causes of acute conflict. *International Security*, 16(2), 76–116.
- Homer Dixon, T. F. (1998). *Environment, scarcity, and violence*. Princeton, NJ: Princeton University Press.
- Homer Dixon, T. F. (2002). *The ingenuity gap: Facing the economic, environmental and other challenges of an increasingly complex and unpredictable future*. New York: Vintage.
- Jacobson, J. (1988). *Environmental refugees: A yardstick of habitability*. Washington, DC: Worldwatch Institute.
- Kakonen, J. (Ed.). (1994). *Green security or militarized environment?* Brookfield, VT: Dartmouth.
- Kaplan, R. (1994). The coming anarchy. *The Atlantic Monthly*, 273(2), 44–76.
- Kelly, K., & Homer Dixon, T. F. (1998). The case of Gaza. In T. Homer Dixon & J. Blitt (Eds.), *Ecoviolence: Links among environment, population, and security* (pp. 67–107). Lanham, MD: Rowman & Littlefield.
- Koremenos, B., Lipson, C., & Snidal, D. (Eds.). (2001). The rational design of international institutions [Special issue]. *International Organization*, 55(4), 761–1103.
- Levy, M. A. (1995). Is the environment a national security issue? *International Security*, 20(2), 35–62.
- Litfin, K. T. (Ed.). (1998). *The greening of sovereignty in world politics*. Cambridge: MIT Press.
- Lonergan, S. C. (2001). Water and conflict: Rhetoric and reality. In P. F. Diehl & N. P. Gleditsch (Eds.), *Environmental conflict* (pp. 109–124). Boulder, CO: Westview Press.

- Lowi, M. R. (1999). Water and conflict in the Middle East and South Asia: Are environmental issues and security issues linked? *Journal of Environment & Development*, 8(4), 376-396.
- Malthus, T. (1798). *An essay on the principle of population*. London: John Murray.
- Matthew, R. A. (1999). Conclusion: Settling contested grounds. In D. H. Deudney & R. A. Matthew (Eds.), *Contested grounds: Security and conflict in the new environmental politics* (pp. 291-302). Albany: SUNY Press.
- Matthews, J. T. (1989). Redefining security. *Foreign Affairs*, 68(2), 162-177.
- Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4), 725-753.
- Percival, V., & Homer Dixon, T. F. (1998). The case of Rwanda. In T. Homer Dixon & J. Blitt (Eds.), *Ecoviolence: Links among environment, population, and security* (pp. 201-222). Lanham, MD: Rowman & Littlefield.
- Percival, V., & Homer Dixon, T. F. (2001). The case of South Africa. In P. F. Diehl & N. P. Gleditsch (Eds.), *Environmental conflict* (pp. 13-35). Boulder, CO: Westview Press.
- Rabe, B. G. (2004). *Statehouse and greenhouse: The emerging politics of American climate change policy*. Washington, DC: Brookings Institution Press.
- Raleigh, C., & Urdal, H. (2007). Climate change, environmental degradation and armed conflict. *Political Geography*, 26(6), 674-694.
- Ronnfeldt, C. (1997). Three generations of environment and security research. *Journal of Peace Research*, 34(4), 473-482.
- Salehyan, I. (2008). From climate change to conflict? No consensus yet. *Journal of Peace Research*, 45(3), 315-326.
- Stern, N. (2006). *Stern review on the economics of climate change*. Retrieved May 12, 2010, from <http://www.hm-treasury.gov.uk/sternreview-index.htm>
- Theisen, O. M. (2008). Blood and soil? Resource scarcity and internal armed conflict revisited. *Journal of Peace Research*, 45(6), 801-818.
- Tir, J., & Diehl, P. F. (2001). Demographic pressure and interstate conflict. In P. F. Diehl & N. P. Gleditsch (Eds.), *Environmental conflict* (pp. 58-83). Boulder, CO: Westview Press.
- Ullman, R. (1983). Redefining security. *International Security*, 8(1), 129-153.
- Urdal, H. (2005). People vs. Malthus: Population pressure, environmental degradation, and armed conflict revisited. *Journal of Peace Research*, 42(4), 417-434.
- Westing, A. H. (1992). Environmental refugees: A growing category of displaced persons. *Environmental Conservation*, 19(3), 201-207.

30

TOTALITARIANISM AND AUTHORITARIANISM

EDWARD WEBB

Dickinson College

In the 20th century, some of the darker consequences of the industrial and political revolutions of the late 18th and early 19th centuries appeared: industrialized slaughter and mass terror organized by powerful states against their own societies. Events such as the Holocaust, Stalin's terror, and China's Cultural Revolution challenged political scientists to explain how and why states could govern in such ways. Although the century ended with a wave of democratization in many parts of the world, different types of nondemocratic regimes that had been pervasive outside western Europe and North America persisted in smaller but still very significant numbers. In the 21st century, political science must continue to analyze nondemocratic regimes and to ask questions that have challenged the discipline for decades (at the very least since the rise of fascism, Nazism, and Stalinism in the period between the two World Wars):

- How can we most usefully categorize and distinguish the types of nondemocratic regimes?
- How do such regimes arise?
- How do they maintain their grip on power?
- Under what circumstances can they fall?

This chapter explains some of the challenges to adequately defining and describing different kinds of nondemocratic regimes and introduces theoretical perspectives and approaches to their study. It then explores some more specific empirical questions on the rise, performance, and demise

of such regimes in the 20th and early 21st centuries. Last, it offers some suggestions for fruitful directions for further research.

Theoretical and Definitional Questions

Definitions and Typology

Although the theoretical origins of modern authoritarianism may be found in classics of political thought including Plato's *Republic*, Hobbes's *Leviathan*, and Rousseau's *On the Social Contract*, most modern political science analysis is informed, at least implicitly, by Max Weber's (1947/1964) concept of *legitimate authority*. According to Weber,

There are three pure types of legitimate authority. The validity of their claims to legitimacy may be based on: 1. Rational grounds—resting on a belief in the “legality” of patterns of normative rules and the right of those elevated to authority under such rules to issue commands (legal authority). 2. Traditional grounds—resting on an established belief in the sanctity of immemorial traditions and the legitimacy of those exercising authority under them (traditional authority); or finally, 3. Charismatic grounds—resting on devotion to the specific and exceptional sanctity, heroism or exemplary character of an individual person, and of the normative patterns of order revealed or ordained by him (charismatic authority). (p. 328)

Modern liberal democracy is based on legal authority, the legitimacy granted by the results of free and fair elections held according to procedures to which, at least notionally, all citizens have given their assent. Other modern regime types draw on a broader array of justifications for their rule. This does not mean that nondemocratic regimes derive their authority exclusively from traditional or charismatic sources, although both kinds of authority have existed and do exist in nondemocratic regimes, such as in some of the monarchies of the Middle East (for all that they are modern creations, their appeal is to traditional allegiances) or the charismatic appeal of a Hitler, a Franco, or a Peron. Some political scientists have adapted Weber's concept of traditional rule to develop modern categories of *neopatrimonialism* or *sultanism* to describe many of the regimes of the Middle East and sub-Saharan Africa, but not all scholars of authoritarianism use these terms.

Most authoritarian regimes rely on a mix of legitimacy and coercion. The tools available to a regime in control of a modern state both to communicate its legitimacy and to apply coercion far outstrip what was available to historical autocrats, even absolutist monarchs such as Louis XIV of France. Modern states are capable of organizing a whole society, should they choose to, through communications technology, pervasive bureaucracy, and sheer firepower.

Authoritarian systems are nondemocratic systems. Put another way, authoritarianism is the absence of or limit on *polyarchy* (Dahl, 1979) or even a limit on politics per se:

In authoritarian systems there is only a limited form of politics, for power struggles among factions in one party regimes and disagreements among soldiers or bureaucrats are not the same as real politics, which must operate in the context of a civic culture. (Pye, 1990, p. 15)

Although the absence of democracy is at the heart of the definition of authoritarianism, it is important not to equate authoritarianism simply with the absence of elections. Elections are used as a tool of legitimation by virtually all regimes, given the almost universal need to claim to rule on behalf of the people. There are many techniques whereby an election or referendum can be manipulated to achieve the outcome desired by the incumbents or to limit the outcome to a circumscribed range of possibilities. In the Islamic Republic of Iran, candidates for public office compete in genuinely competitive elections, but the religious authorities exclude many possible candidates from competing. In Egypt, the Muslim Brotherhood is banned from operating as a political party, meaning that its candidates must run as independents. Moreover, the process of voting itself is often marred by ballot-stuffing, voter intimidation, bribery, and other practices. In single-party regimes, of course, only candidates of the ruling party may run. It is not the case that elections in all these cases are entirely meaningless. What is crucial is that the ruling elite finds means to insulate itself from effective challenge via the ballot box.

What varieties of regime are there? A straight binary division between democracies and nondemocracies is likely to obscure more than it reveals. However, how one develops and applies a typology to regimes that enables fruitful development of theory and empirical analysis is an issue that has yet to achieve a consensus among political scientists and continues to present challenges to theory development. In a recent argument about the problems of regime typology, Stephen Hanson and Jeffrey Kopstein (2005) warn of generating "as many taxonomies as social researchers" (p. 77).

One approach would be to have a finely grained typology that has many separate categories—for example, dictatorship could be divided into *autodictatorship* (no elections), *monodictatorship* (noncompetitive elections), and *semidictatorship* (semicompetitive elections; Brooker, 2000), and then each subtype could be further divided into military, party, or personal rule. Or one might choose a sliding scale from nondemocracy to near-democracy (Diamond, Linz, & Lipset, 1988). This approach is used by organizations such as Freedom House, which assigns a democracy "score" to regimes worldwide. Another approach would have a few broad categories and then qualify the members of those categories with adjectives, an approach designed to "adequately characterize the diverse regimes that have emerged in recent years [while] maintaining conceptual validity by avoiding conceptual stretching" (Collier & Levitsky, 1997, p. 448). This approach yields concepts such as "soft" authoritarianism and "illiberal" democracy (see Chapter 31, "Semi-Authoritarianism," and Chapter 32, "Models of Democracy," in this handbook).

How one categorizes different regimes will be informed by which characteristics one believes to be theoretically important. In a 2007 article, Axel Hadenius and Jan Teorell consider institutional variables the most important indicators of which authoritarian regimes are likely to be sustainable and which prone to democratization. Their typology first sets a firm dividing line between democracies and autocracies and then further divides the latter category into monarchies, military regimes, and electoral regimes, whether no party, single party, or limited multiparty. Bradley Glasser (1995), in contrast, develops a typology of Middle East regimes based on access to resources, in order to explain the durability of some authoritarian regimes in the region and the relative liberalization of others.

Identifying which are the most productive typologies remains part of the challenge for future research. For an excellent overview and discussion of work on typologies to date, see Paul Brooker (2000). It may be helpful here to identify certain broad categories of regime, based on institutional structure, before considering the particular problems raised by the term *totalitarian*.

Party regimes are those in which government is by a sole or dominant party. A *single party regime* has one party identified with the state and permits no others to

operate. Even among this group, there is potentially considerable variation in terms of the ideology through which the party justifies its rule, whether the party has its own security structures or militias, and other variables. A *dominant party regime* allows limited competition from other parties. Party regimes may be of any ideological complexion, although two common types are of the far left (communist or socialist) on one hand and the far right (fascist or nationalist) on the other. *Populist* parties might combine elements usually associated with both left and right. Arguably, ideology is less useful in analyzing these regimes than in recognizing the party as primarily an instrument for maintaining power.

Another common kind of nondemocratic system is the *military regime*, although here again the label covers, theoretically and in practice, a great range of systems. The military may rule directly, through a dictator or ruling council. It may rule indirectly, through alliance with civilian politicians that it chooses and who rule at its pleasure. It may rule through a *front party*, at which point it of course displays elements of the party and military regimes. A military regime may represent the military as a whole or only one section of it. It may aspire to permanent rule or see its intervention into politics as a temporary measure to redress some severe problem, such as a threat to national security or a threat to the interests of the military itself as an institution.

A third kind of nondemocratic regime is the *personalist regime*. This is the dictatorship of an individual and is more likely than the two previous types to be based on traditional or charismatic, rather than legal, authority. One key distinction between this and the previously discussed types is that the basis of the regime's claim to rule is the person of the ruler or rulers, whether they are in power as a result of their own nature or achievements or via descent or other connection to revered authority.

One of the enduring classics on authoritarian regimes is Hannah Arendt's (1951/1973) *The Origins of Totalitarianism*. Yet there has been much debate about whether the term *totalitarianism* has much analytic utility separate from the broader notion of authoritarianism. Written in the aftermath of World War II, Arendt's work was primarily an attempt to explain the cataclysm that had overtaken Europe at the hands of the Nazi regime. How could human progress have become so badly derailed? At least part of the answer was in the unprecedented ability of modern movements and states to dominate society and individuals utterly: "Totalitarianism has discovered a means of dominating and terrorizing human beings from within" (Arendt, 1951/1973, p. 325). Another part was the rise of mass politics; of the rule of the mob and the failure to uphold the rights of minorities and individuals; and of the attempt to impose uniformity on people, to make them into a whole, a total unity.

Another classic arguing for the distinctiveness of totalitarian regimes is Carl Friedrich and Zbigniew Brzezinski's (1965) *Totalitarian Dictatorship and Autocracy*. They

argue that all totalitarian dictatorships possess the following characteristics:

1. An elaborate ideology, consisting of an official body of doctrine covering all vital aspects of man's existence to which everyone living in that society is supposed to adhere. . . . 2. A single mass party typically led by one man, the "dictator," and consisting of a relatively small percentage of the total population. . . . 3. A system of terror, whether physical or psychic, effected through party and secret police control. . . . 4. A technologically conditioned, near complete monopoly of control, in the hands of the party and of the government, of all means of effective mass communication. . . . 5. A similarly technologically conditioned, near complete monopoly of the effective use of all weapons of armed combat. 6. A central control and direction of the entire economy. (p. 22)

They did not present this list of six traits as exhaustive but argued that they had "been generally acknowledged as the features of totalitarian dictatorship" (p. 23).

A totalitarian regime is one that successfully controls all aspects of society, abolishing the distinction between public and private, aspiring even to control the most intimate aspects of individuals' lives and thoughts. Powerful literary portrayals of such regimes can be found in Orwell's *1984*, Aldous Huxley's *Brave New World*, Terry Gilliam's film *Brazil*, and other works. One objection to the application of the term in political science is that arguably such societies have not existed historically, even if regimes have aspired to such levels of control. If there have been totalitarian regimes, at most the term could be applied to the Third Reich, the Soviet Union under Stalin, arguably Iraq under Saddam Hussein, and possibly North Korea. Fascist Italy should not be included:

Fascism had aspirations to be "totalitarian"; Mussolini virtually invented the term. But . . . it is clear that Mussolini's grip on Italian society was not as firm, his influence so pervasive, as that of a Hitler or a Stalin. Fascism left huge areas of Italian life practically untouched. (Lyttelton, 1987, p. 1)

Some question whether even the Hitler or Stalin regimes, with their pervasive propaganda and state terror, did not themselves still leave significant areas of life untouched. Saddam's regime and leader cult aspired to invade all areas of life, to turn children against parents, to make everyone an informer, and succeeded to a remarkable extent (al-Khalil, 1990). And it is hard to make any firm assessment of North Korea since it remains the "hermit kingdom"—a closed society ruled by a paranoid dynasty through many of the tools used by the other named regimes but very difficult for outsiders to investigate and assess (Daniel Gordon's documentary *A State of Mind* is a rare window on the tools used by the secretive regime to manage society).

Some have objected to the term *totalitarianism* as theoretically empty, for instance on grounds that it is simply

a particular instance along the spectrum of authoritarianism (Barber, 1969), while others want to apply it quite widely in recognition of the ambitions of many modern authoritarian regimes, even if they ultimately fall short of total management of society (Friedrich, 1969). But one of many important contributions to the field by Juan Linz (1975) argued for a threefold distinction between *totalitarian*, *authoritarian*, and *democratic* regimes (setting aside anomalies such as sultanistic regimes), arguing that totalitarian and authoritarian were distinctive types of nondemocratic regimes rather than instances along a continuum. For Linz, the characteristics of authoritarian regimes that distinguished them from totalitarian regimes were the presence of limited political pluralism and either demobilization of the population or limited and controlled mobilization.

Theoretical Approaches

Brooker has argued that there can be no theory of authoritarianism *per se*. Rather, we can study different kinds of authoritarian regimes and attempt to account for different aspects of them—how they come to exist, how they maintain power, and what can cause them to fail. Nevertheless, in the attempts made to date to explain these different aspects, we can discern several broad approaches, including those rooted in psychology, in the analysis of ideas and ideologies and in structures and institutions.

One older approach is based in psychology, particularly associated with the Frankfurt School, blending elements of Freudian and Marxist thought. The classic work here is that produced by Theodor Adorno and collaborators (1950), *The Authoritarian Personality*. This approach takes individual characteristics as the enabling condition of the rise of mass movements, explaining those characteristics themselves as produced by the social conditions of modern life. Wilhelm Reich (1970) offers a flavor:

“Fascism” is only the organized political expression of the structure of the average man’s character . . . the basic emotional attitude of the suppressed man of our authoritarian machine civilization and its mechanistic mystical conception of life. It is the mechanistic mystical character of modern man that produces fascist parties, and not vice versa. (p. xiii)

Another approach is interpretive, treating ideas and ideology as having explanatory power. Arendt’s (1951/1973) work, while also sociological and historical, certainly pays attention to this dimension. More recently, Lisa Wedeen (1999) has applied an interpretive lens to the Syrian leader cult, drawing striking conclusions about how its very implausibility and absurdity make it a powerful tool for retarding the emergence of an effective civil society.

The historical-institutional approach, or historical sociology, which ultimately draws from both Marx and Weber, has produced much fruitful work in this domain. Skocpol’s work on social revolutions is a very influential

example of this approach, but there are many others, including Linz and Guillermo O’Donnell. Recent exemplars include the contributors to the special issue of *Comparative Politics* in January 2004 on “Enduring Authoritarianism: Lessons From the Middle East for Comparative Theory” and the work of Jason Brownlee (2007), who studies ruling parties.

Brooker (2000) proposes a forensic approach to the study of the rise, maintenance, and fall of nondemocratic regimes; in other words, their analysis must consider motive, means, and opportunity. This is a broad comparative historical approach that does not commit the researcher *a priori* to favor agency over structure or vice versa, for example, and is designed to allow a broad range of comparison among nondemocratic regimes and between them and their democratic counterparts. It is a useful open framework within which more specific questions can be considered.

Empirical Questions

The Origins of Totalitarian and Authoritarian Regimes

Many modern authoritarian regimes have emerged as the result of revolutions, beginning with the Terror that followed the French Revolution and including the Bolshevik Revolution of 1917 in Russia and the later revolution in China led by Mao Tse-tung. This last established an authoritarian regime that has persisted into the 21st century in the world’s most populous country. The Iranian Revolution from 1977 to 1979 ushered in a unique form of partially democratic rule, the Islamic Republic.

Theda Skocpol (1979) established a structural and comparative historical framework for the analysis of revolutions in *States and Social Revolutions*, which remains required reading in the field. She concluded that the classical Marxist prediction of revolutionary change driven by class conflict did not pay sufficient attention to state power, failing to “adequately explain the autonomous power, for good or ill, of states as administrative and coercive machineries embedded in a militarized international states system” (p. 292). Revolutions largely came about due to the weakening of the state’s ability to maintain its monopoly on the legitimate use of force, frequently due to a combination of international pressures—war, above all—and socioeconomic pressures within. She also argued that in direct contradiction to the Marxist prediction of the withering away of the state after a successful revolution, rather “the new-regime states that emerged in France, Russia, and China alike were stronger and more autonomous within society” (p. 285). While the ideological and socioeconomic complexions of these three postrevolutionary regimes were quite different, the strengthening of the state was a common outcome in all. In the latter two cases, as in many other 20th-century examples of authoritarian

regimes, the state's authority was seen as a necessary tool to apply to the problem of "catching up" with more economically developed countries.

Iran's revolution posed a challenge to this understanding of social revolution, as Skocpol (1982) conceded in an article published not long after her 1979 work. In contrast to her account of those earlier revolutions, she found herself having to recognize a larger independent causal force for ideas and ideology and to stress agency over opportunity. The Iranian revolution came about not in a state weakened by war, but in one where the U.S.-backed regime of the Shah had at its disposal significant military, police, and intelligence resources. This revolution did not simply "come" but was definitely "made" by a coalition of dissatisfied social forces.

Capturing the state and then using its institutions as machines to transform and better society are common goals of revolutions, whatever the details of their mobilizing ideology. Rather than pay attention to ideology, then—to "motive," in Brooker's approach—it is possibly more fruitful to consider the questions of means and opportunity.

Opportunity has often emerged through the breakdown of the existing social and political order, such as via wars, economic disruption, and failure of institutions, as seen in Italy and the Weimar Republic of Germany between the two World Wars. In these circumstances, revolutions can be genuinely *social* in the sense that they are driven by a broad movement of citizens seizing power, although it is usually a rather small group of people who seize the moment to knock out the previous regime. But in the latter half of the 20th century, the dynamic was different because state institutions, and the instruments of state coercion above all, were stronger than in earlier times. Regime change tended to be imposed through military force, from within or without: "Once decolonization was completed, with modern military establishments successfully installed, then social revolutions became much less likely—although military coups of various sorts have become very frequent" (Skocpol, 1979, p. 290). The ability of states to maintain order, to prevent mass mobilization against themselves, has developed, particularly as a result of technological advances. The prospects for social revolutionaries have correspondingly diminished such that some recent analyses more or less exclude any significant role for ordinary citizens in the rise of authoritarian regimes, seeing authoritarianism instead as the outcome of elite decisions, with citizens more or less innocent bystanders who would on the whole prefer democracy (e.g., Bermeo, 2003).

One variation is bureaucratic or elite "revolutions," an idea particularly associated with O'Donnell's (1973) concept of *bureaucratic authoritarianism*, based on his studies of South America and developed and refined in his subsequent work. But this kind of top-down revolution has a longer pedigree, as Ellen Kay Trimberger (1978) found when she studied the Meiji Restoration in 19th-century Japan, the Ataturk regime that founded modern Turkey in the early 20th century, the regime of Gamal Abdel Nasser

in mid-20th-century Egypt, and the Peruvian military junta that took power in 1968. These were examples of what she terms *revolution from above*, to distinguish them from regimes produced either by mass-based revolutionary movements or by straightforward military coups. Revolutions from above, she argues, were revolutionary in that through force or the threat of force a new group took over the central state apparatus and destroyed the economic and political power of the previously dominant social group. But revolutions from above were distinctive in that those taking over were senior military and civilian bureaucrats; that along with little or no mass participation, there was relatively little violence; and that there was little appeal to radical ideology, but rather pragmatic, gradual change. Where Trimberger parts company with earlier thinkers on revolutionary change, such as Marx and Weber, is in making the bureaucracy an actor rather than a tool of other actors (dictators, parties, mass movements) and arguing that while many bureaucracies in industrializing countries may tend to be conservative, some can become forces for change rather than continuity. In this she supports O'Donnell's emphasis on state bureaucrats—military and civilian—as the architects of state-driven social change. The regimes he studied in South America, a region in which many countries had long histories of democracy since independence, often followed the pattern of a military coup's leading to either direct rule by the military or indirect rule through a front political party or organization. Like Turkey, but unlike Egypt, those regimes have now given way to democracies once more, as has South Korea, once also described as a bureaucratic-authoritarian regime.

Diminishing opportunities for social revolution as a result of stronger states may direct us to pay particular attention to the means employed in building nondemocratic regimes. Revolutions are almost always achieved through the application of military force, conventional or unconventional. For example, Thomas Hammond (1975) concluded that from the Bolshevik Revolution onward, Communist regimes came to power only with the aid of significant military force. This is clearest in the case of *exported revolutions*, in which the Soviet Red Army either annexed territories directly to the Soviet Union or installed Communist regimes in other states, many in the immediate aftermath of World War II, but also including the counter-revolutions in East Germany (1953), Hungary (1956), and Czechoslovakia (1968). These are not, of course, the same as the social revolutions analyzed by Skocpol in that they were not the overthrow of a weakened political and economic order by organized social forces from below. But *indigenous takeovers* do more closely resemble the classic social revolutions and also rely on military force. The most common pattern, contrary to classical Marxist predictions of the industrial proletariat as the motor of revolution, is of takeovers starting in the countryside and then spreading to the cities. Examples are Mao's China, Yugoslavia, Albania, and Vietnam (Hammond, 1975). Almost all cases of both indigenous Communist takeovers and exported revolutions

occurred during or soon after international wars that weakened the existing order. In other cases of newly emerging authoritarian regimes, the military coup was the most common pattern, particularly in Latin America and much of postcolonial Africa.

How Totalitarian and Authoritarian Regimes Rule and Survive

In order to understand how nondemocratic regimes maintain their power, we might think the obvious place to start would be with coercion and fear—images of the Soviet gulag or South American death squads or pervasive insecurity generated by secret police forces. We do indeed need to understand the workings of these technologies of power and control, but we must recognize that the range of tools available to modern authoritarian states is broader than those that are directly coercive. And we should understand also that even in regimes that are capable of startling brutality toward their own population, the surest way to stay in power is legitimacy, that is, the consent of the governed.

Although, by definition, nondemocratic regimes do not have the legitimacy that comes as a result of free and fair elections, it is typical and important that they claim to rule in the name of and on behalf of “the people.” Arendt (1951/1973) and others speak of *mass politics*, in contrast to *democratic pluralism*: Authoritarianism often rests on the notion of the citizenry as a single, homogeneous body that finds its expression in the regime, party, or leader. Taken to an extreme, as in the Third Reich, this can lead to programs of expulsion or extermination of the “other.”

How do authoritarian regimes acquire and sustain legitimacy in practice? Sometimes a charismatic leader or one claiming legitimacy through descent or religious sanction may be the source of the regime’s claim to power. But more common is a claim by performance, the regime’s keeping its side of an implicit or explicit bargain—obedience on the part of the population met with the provision of services on the part of the state, whether this be in the form of development, social welfare, or the provision of security. There may also be the provision of what we might think of as psychological goods, such as national pride. We might naturally associate this idea of state provision of benefits in return for obedience with regimes of the left. But one historian of 20th-century fascism points out that regimes of the right also tended to seek legitimacy through performance, by promising to deliver rapid industrialization and national independence without the social dislocations associated with the earlier industrializing states, thus appealing to social conservatism and people’s attachment to order. Adrian Lyttelton (1987) describes Italian fascism in these terms:

[It] aimed at modernization without modernity. In other words, it aimed to appropriate the advantages of technical and economic progress, while rejecting the political, cultural, and

social changes that had been associated with industrialization in Britain and the USA. . . . Fascism started from modernist premises (futurism) and ended by espousing policies of a much more traditionalist kind (ruralism); national socialism started from anti modernist premises (“Blood and Soil”), but ended up by promoting rapid industrialization and urbanization. (pp. 437–438)

Similar appeals to “modernization without modernity” have marked many regimes of the later-developing global South. Malaysia’s former Prime Minister Mahathir Mohamad, among others, is well-known for his espousal of a doctrine of “Asian values,” by which he meant his regime would deliver economic development without political liberalization, which, he argued, was not in keeping with his society’s values and traditions. This approach has had some support from Western economists, politicians, and institutions, although certainly not all (Robison, 1996).

The difficulty with performance-based legitimacy is, of course, that the regime must perform. In the 1950s and early 1960s, there was significant support in political science and related disciplines and in the worlds of policy and politics for the idea that authoritarian regimes could aid in the process of catching up to the more advanced industrialized states. While *modernization theory*, popular in the early part of the cold war, posited that economic and political development moved hand in hand in the direction of greater liberalization, the idea that a more directive, interventionist state could produce more rapid economic development became accepted, not only among authoritarian regimes themselves and their supporters, but also among some development economists and political scientists, particularly those who studied the *Asian Tigers*—states such as Taiwan and South Korea that achieved rapid growth under illiberal, interventionist regimes. However, for every Singapore there was a Zaire, the latter being a classic example of a *predatory state*, where the dictator and his cronies stripped the country of resources and impoverished rather than developed it (the World Bank has termed predatory states *transfer economies*, that is, those in which political power determines the transfer of resources extracted from society to a particular, nonproductive group or class). Many military regimes pinned their legitimacy on outperforming the allegedly inefficient civilian regimes they had replaced, but as early as the 1970s, cross-national analysis already showed that military regimes varied as widely as their civilian counterparts in their performance, that there was no reason to expect that military regimes would deliver more rapid or more effective development than civilian regimes *as a general rule*, even while individual regimes of both kinds might perform well (McKinlay & Cohan, 1975). For arguments about the circumstances under which state intervention is developmental rather than predatory, see in particular the work of Peter Evans (1995).

The planned economies of the Soviet Bloc did achieve rapid industrialization and were able to compete militarily

with the earlier-industrializing West. But ultimately many projects of state socialism built on the Soviet model ended up providing neither bread nor freedom, neither prosperity nor liberty. An American Political Science Association president noted the following in his annual address as Soviet Communism was collapsing:

There is a growing awareness that authoritarian rule, whether Leninist or not, can be a liability in achieving progress. The idea that centralized authority enhances the state's ability to shape society has been dealt a blow by the record of performance of the states that have tried to carry the idea to the extreme. (Pye, 1990, p. 9)

Ideology and cultural authoritarianism are also useful tools for authoritarian regimes. While not even the totalitarian states have achieved the nightmare vision of Orwell's *1984* in terms of constant and ubiquitous surveillance and propaganda, of language stripped of meaning and citizens' very thoughts controlled by the state, most authoritarian regimes have intervened significantly in the domain of culture and public expression, through censorship and propaganda, to mold the environment and channel public discourse in useful directions. A striking instance of this is the *mass spectacle*, whereby regimes command masses of people to perform complex displays, providing both a sign of the regime's power and a disciplining exercise for the performers themselves. The German Marxist critic Siegfried Kracauer discussed the meaning of such displays in his essay collection *The Mass Ornament* (1995), and they can be seen in filmmaker Leni Riefenstahl's propaganda film *Triumph of the Will*, documenting the Nazi rally at Nuremberg in 1934. More recent North Korean and Syrian examples are shown in *A State of Mind* and discussed in Wedeen's (1999) *Ambiguities of Domination*. Other, more day-to-day conditioning of populations can be achieved through mass organizations such as political parties, regime-controlled trade unions, and similar institutions that function to discipline populations but not necessarily to mobilize them (see Kasza, 1995; Unger, 1974).

The tools of fear and coercion are an important complement to the tools of legitimacy and mobilization. Military and paramilitary forces, secret police and intelligence services, prison camps and torture chambers have all been key instruments of 20th-century authoritarian regimes and remain important. The eternal puzzle of *quis custodiet ipsos custodes?*—who guards the guards themselves?—applies here. Saddam Hussein maintained a complex system of rival intelligence agencies, spying on each other as well as the general population, the military, and so forth, and ultimately reporting only to him. The interesting questions are not so much about the functioning of such institutions but about how and why they break down in some circumstances while continuing to perform in others. Why did the Shah's forces stop firing on revolutionary crowds but People's Army troops did

not refuse to crack down on protesters in Tiananmen Square a decade later?

Authoritarian Regimes in an Age of Democratization

Since the collapse of the Soviet Union and its satellite regimes, and the wave of liberalization in Latin America that together make up what Samuel Huntington (1993) described as the "the third wave of democratization," much attention has been paid to transitions from authoritarianism to democracy (see Chapter 33, titled "Processes of Democratization," in this handbook). However, contrary to some who saw in the end of the cold war an end to ideological competition and the beginning of an age in which liberal democracy and free-market capitalism would spread worldwide, political scientists interested in authoritarianism have plenty of material to study in the 21st century. There are aborted or semitransitions, yielding what Fareed Zakaria of the Cable News Network calls "illiberal democracies" or semiauthoritarian regimes, such as Russia. And while there have been very real transitions to democracy in parts of the world, authoritarianism persists in China, North Korea, Central Asia, much of the Middle East, and many countries of sub-Saharan Africa.

Explaining the resistance of the remaining autocracies to global pressures toward liberalization, or, on the other hand, resisting the "end of history" mind-set and analyzing the authoritarian regimes on their own terms rather than as aberrations, has provided interesting challenges for students of comparative politics. Many of those specializing in the Middle East, for example, have been concerned with these questions. Syrian analyst Aziz Al-Azmeh (1994) argues for the importance of ideas and discourse in understanding the persistence of authoritarianism in the Arab world. The discourse of democracy has been co-opted by accident or design in a way that doesn't engage with core democratic ideas or values. Rather, a "populist discourse on democracy" (p. 121) has come to dominate, and populism is an ally of authoritarianism, not least because of the identification of "the people" with the state. Brownlee's (2007) cross-regional analysis takes institutions, rather than discourses, as the basis of its explanation of the resistance of regimes to democratizing pressures. His argument is that in the cases he studies, the key to regime survival is the creation of a dominant party that provides incentives to elites to stay aligned with the regime and can punish those who defect. The absence of such an institution in the case of Iran, for example, means elite politics are more fractious and politics in general more dynamic and open than in Egypt. This institutional approach deliberately compares Middle Eastern cases with those outside the region, rejecting any exceptionalist explanation for its relative lack of liberal democracy. Nevertheless, scholars continue to try to explain the particular problem of the Arab states as "the world's most unfree region" (Schlumberger, 2007, p. 5). Contributors to Oliver Schlumberger's edited volume

consider state–society relations, the structure of the regimes themselves, the interaction of economics and politics, and the interaction of states with the international arena in their search for explanation. The field remains rich in research possibilities.

Directions for Future Research

Areas in which future research will be necessary include, but are not limited to, questions of transitions from authoritarian rule and, conversely, its persistence; state–society relations and the role of civil society under different regime types; civil–military relations; ideology; discourses; propaganda; censorship; and cultural authoritarianism.

Some important issues to explore in the area of transitions and persistence include international pressures and effects. Is there a domino effect for democratization, as U.S. cold warriors once feared regarding communism? Is it possible to impose democracy successfully, as has been attempted in Iraq? What is the role of international institutions and international nongovernmental organizations? Civic culture, or civil society, the presence of public activities, opinion, and organized life outside the control of the state or regime are now considered essential to the healthy functioning of democracy. This has led many donor organizations, governmental and otherwise, to devote considerable resources to building civil society organizations in nondemocratic states, with mixed success, particularly given how easily authoritarian or semiauthoritarian regimes can paint such efforts as sinister “outside interference” (Grodsky, 2007). This is an area in which policy-relevant research is needed.

Economics may also matter a great deal, as it has in 20th-century authoritarianism. What is the relationship between economic liberalization or, conversely, protectionism and political liberalization or authoritarianism? Is it effective to liberalize in one domain but not the other? The question of performance legitimacy in the economic domain persists.

Under the broad heading of state–society relations under authoritarian rule, one key area is civil–military relations, particularly important for societies making a transition away from military rule. Other issues that remain incompletely understood include the effectiveness of propaganda and censorship and the way emerging technologies, particularly in the domain of media and communications, are affecting the ability of regimes to dominate populations and, conversely, the extent to which they might enable the growth of more autonomous and effective civil society organizations.

Conclusion

Nondemocratic regimes persist and offer challenges to political science. They vary widely in institutional structure,

source of legitimacy, ideology, and degree of openness, which makes developing theories of authoritarianism per se difficult. But the core questions of how such regimes arise and stay in power, how they perform, and the circumstances under which they become vulnerable to challenge offer interesting, policy-relevant research pathways susceptible to more than one methodological approach.

References and Further Readings

- Adorno, T. W., Frenkel Brunswik, E., & Levinson, D. J. (1950). *The authoritarian personality*. New York: Harper & Brothers.
- al Azmeh, A. (1994). Populism contra democracy: Recent democratist discourse in the Arab world. In G. Salamé (Ed.), *Democracy without democrats? The renewal of politics in the Muslim world* (pp. 112–129). London: I. B. Tauris.
- al Khalil, S. (1990). *Republic of fear: The inside story of Saddam's Iraq*. New York: Pantheon Books.
- Arendt, H. (1973). *The origins of totalitarianism*. San Diego, CA: Harvest. (Original work published 1951)
- Barber, B. R. (1969). Conceptual foundations of totalitarianism. In C. J. Friedrich, M. Curtis, & B. R. Barber (Eds.), *Totalitarianism in perspective: Three views* (pp. 3–52). New York: Praeger.
- Bermeo, N. (2003). *Ordinary people in extraordinary times: The citizenry and the breakdown of democracy*. Princeton, NJ: Princeton University Press.
- Brooker, P. (2000). *Non democratic regimes: Theory, government, and politics*. New York: Palgrave Macmillan.
- Brownlee, J. (2007). *Authoritarianism in an age of democratization*. Cambridge, UK: Cambridge University Press.
- Collier, D., & Levitsky, S. (1997). Democracy with adjectives: Conceptual innovation in comparative research. *World Politics*, 49(3), 430–451.
- Dahl, R. (1979). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.
- Diamond, L., Linz, J. J., & Lipset, S. M. (Eds.). (1988). *Democracy in developing countries*. Boulder, CO: Lynne Rienner.
- Enduring authoritarianism: Lessons from the Middle East for comparative theory [Special issue]. (2004). *Comparative Politics*, 36(2).
- Evans, P. (1995). *Embedded autonomy: States and industrial transformation*. Princeton, NJ: Princeton University Press.
- Friedrich, C. J. (1969). The evolving theory and practice of totalitarian regimes. In C. J. Friedrich, M. Curtis, & B. R. Barber (Eds.), *Totalitarianism in perspective: Three views* (pp. 123–164). New York: Praeger.
- Friedrich, C. J., & Brzezinski, Z. K. (1965). *Totalitarian dictatorship and autocracy* (2nd ed. rev.). Cambridge, MA: Harvard University Press.
- Glasser, B. L. (1995). External capital and political liberalizations: A typology of Middle Eastern development in the 1980s and 1990s. *Journal of International Affairs*, 49(1), 45–73.
- Gordon, D. (2006). *A state of mind*. New York: Kino on Video.

- Grodsky, B. (2007). Looking for *solidarność* in Central Asia: The role of human rights organizations in political change. *Slavic Review*, 66(3), 442–462.
- Hadenius, A., & Teorell, J. (2007). Pathways from authoritarianism. *Journal of Democracy*, 18(1), 143–156.
- Hammond, T. (1975). A summing up. In T. Hammond (Ed.), *The anatomy of Communist takeovers* (pp. 638–643). New Haven, CT: Yale University Press.
- Hanson, S. E., & Kopstein, J. S. (2005). Regime type and diffusion in comparative politics methodology. *Canadian Journal of Political Science/Revue canadienne de science politique*, 38(1), 69–99.
- Huntington, S. P. (1993). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Huntington, S. P., & Moore, C. H. (1970). Conclusion: Authoritarianism, democracy, and one party politics. In S. P. Huntington & C. H. Moore (Eds.), *Authoritarian politics in modern society* (pp. 509–517). New York: Basic Books.
- Kasza, G. J. (1995). *The conscription society: Administered mass organizations*. New Haven, CT: Yale University Press.
- Kornhauser, W. (1959). *The politics of mass society*. New York: Free Press.
- Kracauer, S. (1995). *The mass ornament: Weimar essays* (T. Y. Levin, Trans.). Cambridge, MA: Harvard University Press.
- Linz, J. (1975). Totalitarian and authoritarian regimes. In F. Greenstein & N. Polsby (Eds.), *Handbook of political science* (Vol. 3, pp. 175–357). Reading, MA: Addison Wesley. (Republished as a stand alone text with a new introduction in 2000 by Lynne Rienner, Boulder, CO.)
- Lyttelton, A. (1987). *The seizure of power: Fascism in Italy 1919–1929* (2nd ed.). Princeton, NJ: Princeton University Press.
- McKinlay, R. D., & Cohan, A. S. (1975). A comparative analysis of the political and economic performance of military and civilian regimes: A cross national aggregate study. *Comparative Politics*, 8(1), 1–30.
- O'Donnell, G. A. (1973). *Modernization and bureaucratic authoritarianism*. Berkeley: Institute of International Studies, University of California.
- Orwell, G. (1949). 1984. London: Secker & Warburg.
- Putnam, R. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Pye, L. W. (1990). Political science and the crisis of authoritarianism. *American Political Science Review*, 84(1), 3–19.
- Reich, W. (1970). *The mass psychology of fascism*. New York: Farrar, Straus & Giroux.
- Robison, R. (1996). The politics of “Asian values.” *Pacific Review*, 9(3), 309–327.
- Schlumberger, O. (2007). Arab authoritarianism. In O. Schlumberger (Ed.), *Debating Arab authoritarianism: Dynamics and durability in nondemocratic regimes* (pp. 1–18). Stanford, CA: Stanford University Press.
- Skocpol, T. (1979). *States and social revolutions: A comparative analysis of France, Russia, and China*. Cambridge, UK: Cambridge University Press.
- Skocpol, T. (1982). Rentier state and Shi'a Islam in the Iranian Revolution. *Theory and Society*, 11(3), 265–283.
- Trimberger, E. K. (1978). *Revolution from above*. New Brunswick, NJ: Transaction Books.
- Unger, A. L. (1974). *The totalitarian party: Party and people in Nazi Germany and Soviet Russia*. Cambridge, UK: Cambridge University Press.
- Weber, M. (1964). The types of authority and imperative coordination. In T. Parsons (Ed.), *The theory of social and economic organization* (A. M. Henderson & T. Parsons, Trans.; pp. 324–423). New York: Free Press. (Original work published 1947)
- Wedeen, L. (1999). *Ambiguities of domination: Politics, rhetoric, and symbols in contemporary Syria*. Chicago: University of Chicago Press.

31

SEMIAUTHORITARIANISM

CHRISTIAN GÖBEL¹*Lund University*

Semiauthoritarianism denotes a form of government that is neither fully democratic nor fully authoritarian. It might be the result of an authoritarian regime's² having adopted some of the features of a democracy, or of a democracy having restricted political or civil liberties. The term is by no means uncontested. As will be seen below, many other terms exist to describe such regimes that fall into the gray zone between democracy and authoritarianism. However, this chapter is not concerned with terminology but rather with the conceptual issues surrounding the phenomenon of countries' not fitting into the existing tripartite typology of democratic, authoritarian, and totalitarian regimes. In this chapter, the notion of a gray zone will be employed to denote this phenomenon. It is suitable for the purposes of this chapter because it does not make any assertions as to the quality of the regimes discussed (i.e., closer to democracy, closer to authoritarianism, stable, in transition). In addition, it provides a good base to discuss the conceptual challenges inherent in classifying regimes. After all, the size of the gray zone depends on the reach of the concepts employed by the researcher. It is not uncharted territory but territory simultaneously charted by two concepts that should not overlap.

Since the 1990s, the number of countries in this gray zone has undoubtedly increased. Although only about 5% of the world's countries were characterized as "partly free" by Freedom House (www.freedomhouse.org) in 1974, their number had grown to as much as one third by 2007.

Although much (but not nearly enough) empirical knowledge exists about each of these cases, there is a dearth of concepts and typologies that would enable us to group them in ways that highlight their similarities and differences. In fact, *semiautocracies*, *hybrid regimes*, *semi democracies*, and related concepts are but a large residual category in which regimes are placed that are neither genuinely democratic nor genuinely autocratic.

There are good reasons for scholars and students of comparative politics to feel uneasy about the large number of countries that do not fit into the existing tripartite regime typology. The fact that so many regimes eschew easy categorization by means of these classes suggests that serious problems exist with respect to the normative and theoretical assumptions and the concepts that underlie this typology. Add to this the considerable methodological difficulties inherent in gathering, systematizing, and making sense of information on so large and so heterogeneous a population of cases, and it is easy to see that this challenge will keep scholars of comparative politics busy for some time to come. This uneasiness extends to the makers of foreign policy in the developed democracies of the Organization for Economic Cooperation and Development. As a consequence of this lack of information, it is difficult for them to assess whether such regimes do or do not pose a threat to national interests, and to decide on the proper instruments to help these regimes develop and become liberal democracies.

The next section explains the roots of the difficulties just outlined, sheds light on efforts made to characterize and bundle the multitude of cases in the gray zone between democracy and authoritarianism, and introduces the most important concepts brought forth since the 1980s. Then, some empirical evidence is summarized to examine how these concepts have emerged from or are applied to such gray-zone regimes. First, a brief analysis of some relevant indices is provided to illustrate the difficulty of drawing the conceptual line between democracies and autocracies. Second, four of the cases most frequently mentioned in the literature on the topic are briefly introduced: Egypt, Malaysia, Russia, and Venezuela.

Thereafter, the policy implications will be discussed. In particular, from an external perspective, the difficulties this lack of knowledge poses for international democracy support are highlighted. Thereafter, possible avenues and conceptual limits of future research are briefly sketched. The last section summarizes the results and rounds off the chapter with a brief conclusion.

Conceptualizing the Gray Zone

The Origins of the Discussion

Only in the past two decades has the issue of regimes that are per definition neither pure democracies nor pure autocracies become a prominent issue on the agenda of comparative political science. This does not mean that such regimes have not existed before. At the outset of the 20th century, there were quite a number of countries in which reasonably free and fair elections among several candidates or even parties were held, but which could not be considered democratic. These regimes tended to follow, as Larry Diamond (2002) puts it, the “optimal path to stable polyarchy, with the rise of political competition preceding the expansion of participation” (p. 23). In the so-called Third Wave of democracy, which began in Portugal in 1974 and swept across most of southern Europe, Latin America, Asia, Africa, and central Europe within two decades, a large number of regimes did not become liberal democracies, however. In a large number of young democracies (most prominently in Africa and Latin America), human rights violations and military interference in politics was so widespread that many observers hesitated to label these regimes “democratic.” After the breakdown of the Soviet Union, the number of regimes that introduced free and fair elections but retained many authoritarian traits continued to grow. Steven Levitsky and Lucan Way (2002) have identified three paths that led to such regimes. The first two originated in a closed authoritarian regime, which either regressed or collapsed and subsequently adopted democratic institutions, often as a result of internal and external pressure. The third path was the regression of a democratic regime. Subsequently, many of these regimes

proved unable to either progress to democracy or reestablish a closed authoritarian regime. At the time, the overwhelming number of democratization scholars held that these regimes were merely transitory and would soon become true democracies. This “transition paradigm” (Carothers, 2002) started to change around the middle of the 1990s and was seriously questioned by the beginning of the 21st century. A growing number of publications suspected that these “hybrid” regimes were not transitional but displayed remarkable stability.

Conceding that there was no inevitable path to full democracy for regimes that either embarked on democratic transition or improved civil liberties raised the question of how these regimes should be conceptualized. Are they democracies or nondemocracies? As easy as this question may sound, it entails two further questions. First, which concept of democracy should form the basis for this research? And even if we could agree on one concept of democracy, where do we draw the border between democratic and non-democratic regimes?

The Conceptual Difficulty of Separating Democracies From Autocracies

Semiauthoritarian regimes are an unwieldy category because they are defined by what they are not: They are not a democracy, nor are they genuinely autocratic. And since autocracies are also defined by what they are not, that is, a democracy, what is left of the equation is that semiauthoritarian regimes are neither democracies nor nondemocracies. At the same time, however, they display elements of both.

The root of this confusing problem lies in how democracy is conceptualized. Scholars do not agree on what procedural characteristics a regime needs to display in order to be called a democracy. The minimalist *electoral democracy* concept, for example, demands elections that are free, fair, inclusive, and meaningful, which not only entails a real chance for the opposition to come to power, but also presupposes a range of civil liberties such as freedom of organization, freedom of speech, and freedom of information. On the other end of the spectrum are maximalist concepts that, in addition to the characteristics just listed, entail a wide range of civil rights, the absence of veto players not legitimized by democratic procedures, horizontal accountability, and the rule of law.

Different as all these concepts may be, they have two features in common that pose considerable difficulties in the process of separating democracies from nondemocracies: (1) They are made up of several criteria, and all of the criteria are necessary elements of a democracy; and (2) more problematically, most of these indicators relate to phenomena that are not either/or conditions but matters of degree. In consequence, the researcher must decide on artificial thresholds that separate existence from nonexistence of the elements inherent in this concept. For example, how

many persons need to be prevented from voting in order for the condition of universal suffrage to be violated? When exactly do elections cease to be free and fair? As a result of the two features above, further conceptual difficulties emerge: Is the half-fulfillment of two conditions equal to the nonfulfillment of one condition? And is a regime that fails on five of eight conditions less democratic than a regime that fails on only one?

The fact that most people would answer the latter question in the affirmative illustrates the biggest problem with existing regime typologies. Per definition, regime typologies must be able to cover all existing cases, and the classes they establish must be mutually exclusive: Each case must correspond to only one type. This means that if the tripartite typology were a true typology, a gray zone could not exist. That it does exist is attributable to the difficulty of establishing a single threshold for a concept that consists of several criteria, each of which is gradual in nature and therefore necessitates the imposition of an artificial threshold.

Attempts to Solve the Gray-Zone Problem

Three main strategies exist to deal with such gray-zone regimes: Conceptualize them as deficient democracies, as subtypes of authoritarian regimes, or as a genuine regime type (hybrid regimes). It needs to be noted that these approaches are not equal to regime types, and the empirical overlaps between them are considerable. This, of course, is again a result of the conceptual difficulties just outlined.

Diminished Subtypes of Democracy

Arguably the most prolific reaction to the conceptual challenges just outlined is the creation of ad hoc concepts to characterize regimes that share most attributes associated with a liberal democracy but not all. In most cases, the deficiency is expressed with an adjective, resulting in terms such as *tutelary democracy*, *illiberal democracy*, *neopatrimonial democracy*, and *delegative democracy*, to just name a few of the literally hundreds of such concepts that have emerged in recent decades. In their famous treatise on conceptual innovations in comparative politics, David Collier and Steven Levitsky (1997) call these concepts “diminished subtypes” of democracy.

Whereas “classic” subtypes share all the attributes associated with democracy and add other features to arrive at increased differentiation (e.g., parliamentary vs. presidential democracies; consensus vs. majoritarian democracies), diminished subtypes are characterized by the lack of one or more of the defining attributes of a liberal democracy. For example, a regime in which horizontal accountability is absent is not a liberal democracy anymore, but might still be more democratic than an electoral democracy. According to one count, the number of such “democracies with adjectives” (Collier & Levitsky, 1997) went into the hundreds, creating confusion and making systematization

difficult. In a notable contribution, a research team associated with Wolfgang Merkel (Merkel, Puhle, Croissant, Eicher, & Thiery, 2003) sought to clear up this confusion. These researchers formulated a very demanding concept of an *embedded democracy* that consists of five “partial regimes,” that is, elections, public participation, effective power to govern, horizontal accountability, and civil liberties. Merkel et al. (2003) also developed 34 indicators to account for “defects” in these partial regimes, with one type of *defective democracy* corresponding to each partial regime. In this way, they identified four basic diminished subtypes of democracy, namely, *exclusive democracy* (which covers any violation of democratic elections and political participation and therefore combines two partial regimes), *tutelary democracy*, *delegative democracy*, and *illiberal democracy*. In other words, the strategy of extending the root concept of liberal democracy into the gray area by subtracting attributes of the root concept persists. Thereby, the root concept is extended toward or even beyond the minimalist concept of an electoral democracy, and a fluid conceptual boundary encompassing both the minimalist and the maximalist concepts is imposed between democracies and authoritarian regimes, and various “defects” or “deficits” mark the difference between electoral and liberal democracy.

Electoral Authoritarianism

“Electoral authoritarianism” is another way of conceptualizing “institutionalized ambiguity” (Schedler, 2002, 2006b). It follows the suggestion of Juan Linz (2000) not to build subtypes of democracy but rather to rely on subtypes of authoritarianism when charting the gray zone. Andreas Schedler maintains that electoral authoritarianism is different from regime hybridity (see below) because regimes in the former category are “neither democratic nor democratizing but plainly authoritarian, albeit in ways that depart from the forms of authoritarian rule as we know it” (Schedler, 2006b, p. 5). They are different from both electoral democracies, which conduct free and fair elections, and *closed autocracies*, which do not stage multiparty elections. Again, however, it is difficult to draw the line between these regime types, with the result that some of the examples that Schedler gives for electoral authoritarian regimes are also textbook examples in some of the literature on hybrid regimes, notably Russia, Egypt, and Malaysia. The proponents of the electoral authoritarianism concept share the stance that these regimes are semiautocratic by their own choice and not because they lack the capabilities to manage the full transition to democracy. On a conceptual level, these authors shed light on the interplay between political and civil rights, and a distinction is made between *hegemonic* and *competitive* electoral authoritarianism. The former pertains to regimes whose “elections and other ‘democratic’ institutions are largely façades, yet they may provide some space for political opposition, independent media, and social organizations that do not seriously criticize or challenge the regime” (Diamond, 2002, p. 26). In

competitive authoritarian regimes, which Levitsky and Way (2002) categorize as hybrid regimes (see below), “formal democratic institutions are widely viewed as the principal means of obtaining and exercising political authority. Incumbents violate those rules so often and to such an extent, however, that the regime fails to meet conventional minimum standards for democracy” (p. 52). Hence, elections are free but hardly fair, with incumbents seeking to control the outcomes by means such as intimidating opposition candidates, corrupting journalists, vote buying, and ballot-box stuffing.

Hybrid Regimes

Instead of extending the range of cases that the concepts *democracy* and *authoritarianism* cover by including (diminished) subtypes of either regime type, Terry Karl (1995, 2005) proposed to treat hybrid regimes as a unique regime type that combines free and fair elections—but not the other traits of a minimalist concept of democracy—with autocratic elements. In fact, the various conceptualizations of hybrid regimes discussed subsequently tend to be based on the two procedural dimensions of democracy identified by Robert Dahl (1971): contestation for political office and public participation in politics. While hybrid regimes tend to grant many or even most of the political rights necessary for political contestation, they tend to control day-to-day politics by systematically excluding certain social groups; by restricting the rights to participate in politics; or by disempowering parliaments, the judiciary, and other control institutions. Often, a relaxation of political rights leads to a tightening of civil liberties, and vice versa.

However, elections are not the only realm in which contestation in *competitive authoritarianism* is played out. As Levitsky and Way (2002) point out, three more arenas are involved. First, in the legislative arena, the opposition can use the legislature as a platform to organize and voice dissent even if it is unable to seize political power. Second, in the judicial arena, judges are often formally independent and can challenge incumbents by declaring a referendum illegal, sanctioning electoral fraud, or protecting the media. Naturally, the incumbents will seek to co-opt or intimidate judges, but such action can be costly in terms of regime legitimacy. Third, the media in competitive authoritarian regimes often enjoy limited or even complete freedom to report on controversial issues and can serve as watchdogs exposing government malfeasance or as mouthpieces of the opposition. As is the case with the judicial arena, the incumbents tend to apply more subtle means than do closed authoritarian regimes for suppressing media freedom.

Marina Ottaway (2003) has added a further important arena, namely, that of gaining public support. The most obvious way of doing so is generating *output legitimacy* by means of providing public services or stimulating economic growth. Where this option is not available (most of the hybrid regimes are developing countries), leaders can rely on personal charisma, co-opt potential opposition by

means of patronage networks, or portray themselves as guarantors of stability, security, and the national interest. A number of later studies confirm that public opinion is indeed an arena that should not be neglected in the analysis of semiauthoritarian regimes.

In a broad manner and almost by intuition, hybrid regimes are often conceptualized as nearer either the autocratic end of the spectrum of political regimes (semiauthoritarian regimes) or the democratic end (semidemocracies). Although such a denomination might be the result of a more or less value-free analytical assessment of the number of democratic preconditions fulfilled or not fulfilled, some authors introduce a normative element to justify their labels. They point out that democracy has today become the only legitimate form of government and that international pressure on nondemocracies to become liberal democracies is rising.³ Because of this, they argue, many authoritarian regimes introduce democratic features (such as elections) to mask their authoritarian character. Hence, Diamond (2002) claims that “virtually all hybrid regimes in the world today are quite deliberately pseudodemocratic” (p. 24). In a similar vein, Ottaway (2003) holds that a defining feature of semiauthoritarian regimes is that “they are carefully constructed and maintained alternative systems” (p. 7). According to Ottaway, such regimes need to be distinguished from countries that are still in transition to democracy and those that strove to become democratic but failed. Although theoretically relevant, it is difficult to ascertain whether a regime is in a protracted transition, has failed to democratize, or is semiauthoritarian by choice. With respect to semiauthoritarian regimes, Ottaway proposes a subdivision into three types of semiauthoritarian regimes. “Semiauthoritarian regimes in equilibrium,” for which she cites Egypt as a textbook case, “have established a balance among competing forces” (p. 20) and are therefore able to persist for a long time. “Semiauthoritarianism in decay” (p. 20) denotes regimes that are regressing toward full authoritarianism, such as Azerbaijan, or from democracy to semiauthoritarianism, such as Venezuela, and that tend to be much more unstable than semiauthoritarian regimes in equilibrium. “Semiauthoritarian states undergoing dynamic change” (p. 20) form a third category and comprise regimes verging on democratization. Ottaway’s examples are Senegal and Croatia, although Senegal was showing, to apply Ottaway’s terminology, signs of “decay” when this chapter was written.

As can be seen from the above discussion, the term *semiauthoritarian* is frequently used as a residual category to cover diminished subtypes of democracy and electoral authoritarian regimes as well. Only at the end of the first decade of the 21st century did the more systematic study of hybrid regimes as a distinct regime type pick up. For example, building on the competitive authoritarianism concept formulated by Levitsky and Way (2002), Joakim Ekman (2009) created a profile of hybrid regimes based on six criteria, namely competitive elections (a minimum condition to be fulfilled in order for a regime to be classified as hybrid), significant levels of corruption, lack of democratic quality, a

problematic press freedom situation, a poor civil liberties situation, and lack of the rule of law. According to Ekman, the more of these criteria that are met, the closer the regime verges on authoritarianism.

In a similar fashion, Heidrun Zinecker distinguishes democracies, autocracies, and hybrid regimes by means of five “partial regimes,” that is, civil rule, polyarchy, rule of law, civilizedness, and political exclusion/inclusion. If none of these conditions is fulfilled, the regime is authoritarian, and if all are fulfilled, it is a democracy. Hybrid regimes are characterized by the presence of at least civil rule and polyarchy, and the nonfulfillment of at least one of the three other conditions. Of particular importance for Zinecker is the exclusion/inclusion dimension, for it functions as a link between regime characteristics and issues of socioeconomic transformation. According to Zinecker, the democratization of a hybrid regime requires the transformation not only of political but also of socioeconomic institutions.

Finally, Leonardo Morlino (2009) straightforwardly argues that conceptualizing hybrid regimes as a genuine regime type can be justified only if it can be proven that regime hybridity is not a transitory phenomenon. For him, only those countries that have been “partly free” in the Freedom House survey for at least 10 years qualify as hybrids. In a way similar to that of Zinecker (2009), Morlino classifies these regimes on the basis of “seven ambits that are important when analyzing any political regime” (p. 290), that is, rule of law, electoral process, functioning of government, political pluralism and participation, freedom of expression and beliefs, freedom of association and organization, and personal autonomy and individual freedom, and arrives at three types of hybrid regimes: *quasi democracies*, *limited democracies*, and *democracies without state*.

Advantages and Disadvantages

Each of the approaches outlined above has advantages and disadvantages. The advantages of creating diminished subtypes are that the perceived deficiencies of individual regimes are highlighted and that democracy can continue to serve as the root concept if a regime is perceived to verge closer to democracy than to authoritarianism. However, there are several notable disadvantages in this strategy. First, some scholars argue that it is unethical to classify one third of the countries in this world by what political scientists perceive to be not their special traits, but their deficits. The more serious disadvantage, however, is the conceptual dilemma this strategy poses. In terms of the strict demands of a typology, since these “diminished subtypes” do not possess all the necessary attributes prescribed for a regime to be called a liberal democracy, they should not be called democratic if liberal democracy is the root concept. In this case, it might be better to take electoral democracy as the root concept and create categorical subtypes.

The concept of electoral authoritarianism seems to solve both these problems. It represents a categorical subtype of authoritarianism that is juxtaposed to a minimalist concept of an electoral democracy. However, it does not address the problem of how to classify those regimes that are democracies but not liberal democracies. Clearly, the strategy of creating diminished subtypes of democracy cannot be applied anymore if the root concept is not a liberal but an electoral democracy. Hence, nonliberal democracies cannot be conceptualized anymore on the basis of the deficits ascribed to them, and entirely new subtypes building on positive or neutral characteristics would have to be created. This is more difficult than it seems.

Treating “hybrid regimes” as a distinct regime type does not solve, but rather dodges, this problem. The concepts presented by Ekman (2009) and Zinecker (2009) allow researchers to broadly measure a hybrid regime’s relative distance from democracy and authoritarianism, and they allow scholars to cluster hybrid regimes around different dimensions that, however, again point to deficits. The more explicit classification proposed by Morlino exemplifies this dilemma. His typology again yields “democracies with adjectives” (Collier & Levitsky, 1997).

Empirical Evidence

Quantitative Evidence

As the previous sections should have made clear, the number of gray-zone regimes depends on the democracy concept applied as well as its operationalization. This pertains to electoral authoritarianism as well because, as pointed out above, authoritarianism is defined as the absence of democracy. This section compares several assessments regarding the number of “hybrid regimes.”

Based on the Freedom House index, which is one of the most popular indices for classifying political regimes, Diamond (2002) counts 73 liberal democracies, 31 electoral democracies, 17 ambiguous regimes, 21 competitive authoritarian regimes, 25 hegemonic electoral authoritarian regimes, and 25 closed authoritarian regimes for the year 2001. If we count as hybrid regimes those that are neither liberal or electoral democracies nor closed authoritarian regimes, the number comes to 62 (far more than one third out of a sample of 167 countries). If hegemonic authoritarian regimes are excluded, almost one fourth (38 regimes) can be classified as hybrids. This tallies well with the Economist Intelligence Unit’s Democracy Index for 2008, which lists 36 out of 167 regimes covered as “hybrid” (p. 2), and Morlino’s findings, which identify 35 hybrid regimes. According to Diamond, the largest percentage of hybrid regimes is found in sub-Saharan Africa, where 27 (17) out of 48 regimes are neither liberal nor electoral democracies nor closed authoritarian regimes (the number in parentheses, in this sentence and below, is the number of hybrid regimes when hegemonic electoral

authoritarian regimes are subtracted). In the Middle East and North Africa, 10 (4) out of 19 fall into this category, 12 (7) out of 27 in the post-Communist regimes, 7 (3) out of 25 in Asia, and 6 (5) out of 33 in Latin America and the Caribbean. According to Ekman's index, 30 countries fulfill at least the minimum criterion of competitive elections and can therefore be classified as hybrid regimes. Russia and Venezuela are positive on all indicators and are therefore the hybrid regimes closest to authoritarianism, whereas Romania and Croatia are positive on only one indicator and are therefore closest to democracy. Although the number of hybrid regimes in all the indices mentioned so far does not differ much if hegemonic authoritarian regimes are excluded from Diamond's typology, there is no consensus on which regimes qualify as hybrid and which do not. A comparison of the indices just introduced yields great variation.

Case Studies

Although, as pointed out above, gray-zone regimes are very different from each other, the different combinations of democratic and authoritarian traits are illustrated in this section by means of four brief case studies. This analysis follows the "competitive authoritarianism" concept and examines the four arenas identified by Levitsky and Way (2002), augmented by the public support dimension identified by Ottaway (2003). The emergence and traits of competitive authoritarianism are outlined for Egypt, Malaysia, Russia, and Venezuela because these four cases figure prominently in the debate, they represent four of six continents, and they display remarkable differences with respect to their "hybridity." As a preface to the case studies, the next two paragraphs briefly shed light on the history of regime hybridity in each case.⁴

Egypt and Malaysia can both be characterized, in Ottaway's terms, as "institutionalized" semiauthoritarian regimes. The institutional base of semiauthoritarianism in Egypt was built under Anwar el-Sadat in the 1970s. He replaced the military regime of his assassinated predecessor, Gamal Abdel Nasser, with a presidential system that nominally guaranteed political and civil rights. In Malaysia, a race riot in 1969 ended the democracy inherited from the British colonizers after the country's independence in 1957. After an 18-month state of emergency, the ruling United Malays National Organization reorganized itself, and gradually a regime combining authoritarian and democratic elements was set up. One crucial difference between the two countries is that in Malaysia, leadership succession took place under the auspices of the former leader, whereas leaders in Egypt remained in power as long as they were alive.

In contrast to these long-standing regimes, Russia's and Venezuela's semiauthoritarianism emerged only in the late 1990s. Again, there are crucial differences between these two cases. In the Russian Federation, semiauthoritarianism gradually replaced an authoritarian regime after

the collapse of the Soviet Union. Semiauthoritarianism was most visibly engineered under Vladimir Putin, who had succeeded the ailing Boris Yeltsin as president of the Russian Federation in late 1999. Whereas Putin arguably had an incumbent advantage in the 2000 presidential elections, the former Venezuelan army colonel Hugo Chavez had not been in office when he won the 1998 presidential elections in Venezuela. However, he was well-known for his leading role in a 1992 coup attempt.

Of the four regimes described here, only Venezuela had the status of an "electoral democracy" in the Freedom House survey, which it lost, however, in 2009 (<http://www.freedomhouse.org/uploads/fiw09/FIW09Tables&GraphsForWeb.pdf>). Until then, opposition parties and candidates were allowed to organize and compete in presidential and parliamentary elections. They had a real chance of winning, but the playing field was not level. As a result of the media restrictions discussed below, Chavez and his followers receive better and more benevolent media coverage, and government resources are used to shore up support in election campaigns. In addition, Chavez's staffing of the National Electoral Council and the use of fingerprint identification equipment have shored up allegations of irregular election procedures because they allegedly allow authorities to connect voters to their ballots. In the other three cases, the opposition is similarly disadvantaged in media coverage, resource endowment, and election procedures and faces additional difficulties. In Malaysia, the electoral system favors the ruling Barisan Nasional. In addition, electoral rolls are frequently manipulated, districts are gerrymandered, and party organization and campaigning are restricted. Nevertheless, opposition parties have been able to win the gubernatorial elections in individual provinces. In Egypt, strict party registration rules, high thresholds for presidential candidates, and electoral manipulation have made it virtually impossible for opposition candidates to contest the presidency. The same is true of Russia, where state-controlled media, high thresholds for candidates, and intransparent elections have made elections increasingly less competitive.

The independence of the judiciary is severely limited in all four countries. In Egypt, the Ministry of Justice controls promotions and compensations for judges, giving the executive undue influence over the judiciary. In addition, exceptional courts appointed by the president try security cases, and the Emergency Law allows the arrest of political activists. Corruption and executive interference hamper judicial independence in Malaysia, Venezuela, and Russia. However, the courts are not entirely dysfunctional in any of the cases.

Freedom of information in Egypt is severely hampered by government ownership of all terrestrial television stations, the licensing requirements for newspapers, and a state monopoly on the printing and distribution of newspapers. In addition, criticism of state authorities and organizations is discouraged by means of defamation suits. The

Russian and Malaysian governments apply similar tactics. Although not unheard of in Russia and Egypt, intimidation is the main form of influencing journalists in Venezuela. However, with the Law on Social Responsibility of Radio and Television, passed in 2004, and the refusal to renew the licenses of critical media, the Chavez administration has started to resort to censorship as well.

As for public support, it is important to note that with the possible exception of Egypt, the strongmen in all four cases have enjoyed genuine popular support. In Venezuela, this support recently started to dwindle, which prompted Chavez to further restrict civil liberties. Similar processes are at work in Malaysia, where regulations banning Muslim conversions to Christianity fueled popular unrest, which was in turn answered with arrests and a media crackdown. Putin has not yet had to face such challenges.

Implications for Democracy Assistance

As the chapter so far has shown, semiauthoritarian regimes can endure for a long time. Given that in such regimes, persons are more important than institutions, it is extremely difficult to predict how a semiauthoritarian regime will react to internal or external crises such as severe economic downturns or the death of a paramount leader. Such a crisis might well lead rival elite groups into a struggle for succession, with potentially negative impacts on the well-being of the population and regional stability. For these reasons and the normative conviction that democracy is the best of all regime forms, governments and civil society organizations in Western industrialized states seek to assist democratization in such countries.

As the above sections should have made clear, however, the study of gray-zone regimes is still in its infancy. Considerable headway is being made with regard to concept formation and regime classification, but a theory of hybrid regimes that could guide democracy assistance does not yet exist.⁵ For this reason, the relationship between democracy assistance and semiauthoritarian regimes is characterized by three major paradoxes that are very difficult to resolve. First, as Ottaway (2003) points out, democracy assistance is one of the major reasons that such a large number of semiauthoritarian regimes emerged in the first place. Development aid and inclusion in international regimes is frequently tied to demands for democratization, which has prompted many authoritarian leaders not to commit to democratization, but to set up democratic façades. Second, democracy assisters demand that semiauthoritarian rulers become accountable to their people but at the same time require these rulers to be responsive to the donor community. More emphasis is placed on the implementation of externally prescribed best practices than on generating homemade solutions in often messy democratic politics. This, one must add, is especially true for those regimes that are dependent on outside economic assistance. Third, given the lack of knowledge about such regimes and the resulting lack of systematic strategy on the side of the donor

community, democracy assisters can never be sure whether their well-meant interference will contribute toward democratizing the regime in question, have no impact at all, or even help extend authoritarian rule.

To overcome these paradoxes, democracy promoters face two major decisions. The first is whether the regime should be changed by means of exerting pressure on the incumbents, propping up the opposition, and (re)designing legal, administrative, and representative structures. The underlying logic is that even if not much is known about the causes for the existence of such regimes, at least the symptoms of authoritarianism can be addressed. Although this approach has the potential to change a regime, it also necessitates the commitment of substantive funds over a long time and runs the great danger of fomenting disorder and even civil war if the intervention fails. The alternative would be less invasive attempts to improve existing structures. This approach might help improve organizational and administrative capacities and the level of education, yet these often are not what the opposition lacks in a country where leaders are simply reluctant to give up power.

The second decision closely relates to the first one and pertains to the instruments and targets of democracy promotion. Again, the lack of knowledge about hybrid regimes makes it difficult to assess the impact of particular instruments on the regime in question. According to Ottaway, global democracy promotion not only tends to avail itself of the same terminology and narratives concerning the benefits of democracy but also tends to prescribe very similar recipes in very different contexts. Among these recipes are election monitoring, propping up civil society organizations, training journalists, and educating the population on the virtues of democracy. However, these strategies, which aim at thwarting the “games semiauthoritarian regimes play” (Ottaway, 2003, p. 137) in the realms of electoral, legislative, and judicial procedures, in the media, and in public support, have proven to be of only limited usefulness. As Zinecker suggests, what are also needed are attempts to remedy the structural factors that brought such regimes into power in the first place.

For example, democracy is very difficult to establish in countries with societies polarized along ethnic or income lines, those with low income standards, those still in the process of state formation, and those where a powerful opposition against the ruling regime does not exist. In addition, the rulers, the opposition, and the general population must be convinced that democracy is indeed the regime form best suited to overcome these difficulties. This, according to Ottaway, necessitates linking politics and policies, which means that democracy assistance should refrain from prescribing outcomes in the form of best practices while at the same time demanding that democratic procedures are followed.

A precondition for establishing a democracy is that rulers can be either convinced or coerced to refrain from manipulating political outcomes, which in turn necessitates either a reformist leadership or a strong opposition willing to follow democratic procedures. In such cases, the chances

for success are best if demands for support come from the regime and are not imposed from the outside. In general, the tools applied to promote democracy must be applied in a context-sensitive manner, which requires intimate knowledge of the case in question. Of general relevance are issues such as whether semiauthoritarianism has persisted for a long time and is likely to persist, the relationship between incumbents and opposition, the (non)existence of likely collaborators in civil society, the regime's (in)dependence on outside funding, the socioeconomic situation, the degree of social cohesion, the quality of existing regime structures, the level and features of corruption, experience with democratic procedures, and so on. This list shows that there are no easy solutions, and it is here that improved knowledge about different types of hybrid regimes and their internal dynamics is clearly necessary in order to design better policies to deal with the challenges emerging from such regimes.

Directions for Future Research

As the preceding sections have shown, the study of gray-zone regimes is still a young and evolving discipline. Future research will have to address three broad issues: conceptual, contextual, and causal. With respect to the first issue, four major roads can be imagined to tackle the dilemma of having to impose an artificial threshold on phenomena that are matters of degree rather than type. First, conceptualize regimes exclusively in degrees, which, however, would in its final consequence mean that the three-part typology in its present form would have to be given up and that one or more gradual concepts extending across all regime forms would have to be found. Second, conceptualize regimes exclusively in types, which would necessitate the creation of a multitude of indicators to capture not only the basic regime types but also their manifold variations. This would be the highly impractical political science equivalent of squaring the circle. Third, further pursue the practice of establishing an artificial threshold between democracies and autocracies while conceptualizing each regime type's quality as a continuum. As seen above, the drawback of this strategy is that it works only with minimalist concepts. The more demanding the concepts become, the larger the gray zone in which they overlap becomes. Thus, most promising is the practice of applying a minimalist concept of an "electoral democracy" and carefully separating it from "electoral autocracies." Fourth, further pursue the third strategy and deal with the gray zone by conceptualizing it as a distinct regime type.

With respect to either strategy, another dilemma has to be addressed when applying these concepts to concrete cases. On one hand, both small-*N* and large-*N* comparative studies have to be based on uniform criteria that allow the researcher to make inferences that transcend these cases. On the other hand, such analysis must take heed of the problem that in different settings, different forms of violating democratic principles might be applied. For example,

whether an election is free and fair has to be decided from case to case because the methods used to influence elections differ and cannot be specified a priori. Hence, the study of hybrid regimes needs to be conceptually and methodologically sound and contextually informed.

A final desideratum for further research concerns questions of causes and effects. This applies especially if the gray zone is classified as a distinct regime type. For example, in what ways do hybrid regimes come about, and how does this history influence their stability? What role do domestic factors such as political culture, the socioeconomic background, and the character of the regime play? Are certain types of hybrid regimes more stable than others? Does a specific regional environment promote the formation of hybrid regimes? And are hybrid regimes harmful to regional stability? In sum, what needs to be developed in the years to come are theories that address these questions.

Conclusion

Semiauthoritarianism represents a true challenge for comparative politics because theories that explain the genesis and impacts of semiauthoritarian or hybrid regimes on regional and domestic stability do not yet exist. The vast amount of work on the subject is still conceptual in nature. This is one of the reasons that international support for democracy faces such great difficulties in picking the right instruments. Comparative politics has begun only very recently to systematically conduct research on such regimes, which truly makes semiauthoritarianism a topic for political science in the 21st century. The conceptual foundation of the existing tripartite regime typology is the root of the difficulties inherent in such theory building, and three different strategies to overcome these difficulties have been introduced in this chapter. In addition, a number of indices have been examined to gauge the number of such regimes existing in the world. Finally, Egypt, Malaysia, Russia, and Venezuela were compared to tease out the similarities in the ways they limit elections, stifle parliamentary and judicial activity, and constrain press freedom, but nevertheless gain a modicum of support. However, these regimes also illustrate important differences with regard to their historical origin, their duration, and the instruments their leaders utilize to stay in power. Future research can pursue a number of promising paths, and this chapter has aimed to enable the reader to gain a better understanding of the challenges involved in theorizing semiauthoritarian regimes. It is hoped that the theoretical and practical relevance of conducting such research has become obvious.

Notes

1. I am indebted to David Kühn and Julia Leininger for their useful comments and suggestions.
2. The term *regime* denotes "the set of government institutions and of norms that are either formalized or are informally

recognized as existing in a given territory and with respect to a given population" (Morlino, 2009, p. 276).

3. At the time of writing, a new line of scholarship has become devoted to exploring whether powerful authoritarian regimes such as China and Russia actively contribute to establishing authoritarianism as a legitimate alternative to democracy.

4. The analysis is based on the Freedom House Country reports for the year 2008. Additional information on Malaysia was obtained from Case (2002).

5. A partial exception is Zinecker (2009), who suggests that hybrid regimes where rent economies predominate can be democratized by the marketization of the economy.

References and Further Readings

- Alvarez, M., & Cheibub, J. A. (1996). Classifying political regimes. *Studies in Comparative International Development*, 31(2), 3–36.
- Bendel, P., Croissant, A., & Rüb, F. (Eds.). (2002). *Zwischen Demokratie und Diktatur: Zur Konzeption und Empirie demokratischer Grauzonen* [Between dictatorship and democracy: Concepts and empirical findings on democratic gray zones]. Opladen, Germany: Leske + Budrich.
- Bogaards, M. (2009). How to classify hybrid regimes? Defective democracy and electoral authoritarianism. *Democratization*, 16(2), 399–423.
- Carothers, T. (2002). The end of the transition paradigm. *Journal of Democracy*, 13(1), 5–21.
- Case, W. (2001). Malaysia's resilient pseudodemocracy. *Journal of Democracy*, 12(1), 43–57.
- Case, W. (2002). *Politics in Southeast Asia: Democracy or less*. Richmond, Surrey, UK: Curzon.
- Casper, G. (1995). *Fragile democracies: The legacies of authoritarian rule*. Pittsburgh, PA: University of Pittsburgh Press.
- Collier, D., & Levitsky, S. (1997). Democracy with adjectives. *World Politics*, 49(3), 430–451.
- Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.
- Diamond, L. (2002). Thinking about hybrid regimes. *Journal of Democracy*, 13(2), 21–35.
- Economist Intelligence Unit. (2008). *The Economist Intelligence Unit's democracy index for 2008*. Retrieved May 30, 2010, from <http://graphics.eiu.com/PDF/Democracy%20Index%202008.pdf>
- Ekman, J. (2009). Political participation and regime stability: A framework for analyzing hybrid regimes. *International Political Science Review*, 30(1), 7–31.
- Hermet, G., Rose, R., & Rouquié, A. (1978). *Elections without choice*. London: Macmillan.
- Howard, M. M., & Roessler, P. G. (2006). Liberalizing electoral outcomes in competitive authoritarian regimes. *American Journal of Political Science*, 50(2), 365–381.
- Jayasuriya, K., & Rodan, G. (2007). New trajectories for political regimes in Southeast Asia. *Democratization*, 14(5), 767–772.
- Karl, T. L. (1995). The hybrid regimes of Central America. *Journal of Democracy*, 6(3), 72–86.
- Karl, T. L. (2005). *From democracy to democratization and back: Before transitions from authoritarian rule* (Working Paper 34). Stanford, CA: Center on Democracy, Development, and the Rule of Law.
- Levitsky, S., & Way, L. A. (2002). The rise of competitive authoritarianism. *Journal of Democracy*, 13(2), 51–65.
- Linz, J. J. (2000). *Totalitarian and authoritarian regimes*. Boulder, CO: Lynne Rienner.
- Merkel, W. (2004). Embedded and defective democracies. *Democratization*, 11(5), 33–58.
- Merkel, W., Puhle, H. J., Croissant, A., Eicher, C., & Thiery, P. (Eds.). (2003). *Defekte Demokratie: Band 1. Theorie* [Defective democracies: Vol. 1. Theory]. Opladen, Germany: Leske + Budrich.
- Mitin, D. (2007). Economic autonomy and democracy: Hybrid regimes in Russia and Kyrgyzstan. *Comparative Political Studies*, 40(12), 1563–1566.
- Morlino, L. (2009). Are there hybrid regimes? Or are they just an optical illusion? *European Political Science Review*, 1(2), 273–296.
- O'Donnell, G. (1994). Delegative democracy. *Journal of Democracy*, 5(1), 55–69.
- Ottaway, M. (2003). *Democracy challenged: The rise of semi authoritarianism*. Washington, DC: Carnegie Endowment for International Peace.
- Rodan, G., & Jayasuriya, K. (2009). Capitalist development, regime transitions and new forms of authoritarianism in Asia. *Pacific Review*, 22(1), 23–47.
- Schedler, A. (2002). The menu of manipulation. *Journal of Democracy*, 13(2), 36–50.
- Schedler, A. (Ed.). (2006a). *Electoral authoritarianism: The dynamics of unfree competition*. Boulder, CO: Lynne Rienner.
- Schedler, A. (2006b). The logic of electoral authoritarianism. In A. Schedler (Ed.), *Electoral authoritarianism: The dynamics of unfree competition* (pp. 1–23). Boulder, CO: Lynne Rienner.
- Schlumberger, O. (2009). Assessing political regimes: What typologies and measurements tell us—and what they don't. In E. Kienle (Ed.), *Democracy building and democracy erosion: Political change north and south of the Mediterranean* (Chapter 2). London: Saqi Books.
- Snyder, R. (2006). Beyond electoral authoritarianism: The spectrum of nondemocratic regimes. In A. Schedler (Ed.), *Electoral authoritarianism: The dynamics of unfree competition* (pp. 219–231). Boulder, CO: Lynne Rienner.
- van de Walle, N. (2002). Africa's range of regimes. *Journal of Democracy*, 13(2), 66–80.
- Zakaria, F. (1997). The rise of illiberal democracy. *Foreign Affairs*, 76(6), 22–43.
- Zinecker, H. (2009). Regime hybridity in developing countries: Achievements and limitations of new research on transitions. *International Studies Review*, 11(2), 302–331.

32

MODELS OF DEMOCRACY

JOHN T. ISHIYAMA, TATYANA KELMAN, AND ANNA PECHENINA

University of North Texas

This chapter examines the various ways in which democracy has been conceptualized: in other words, *models of democracy*. Although the term *democracy* has often been used in the literature, there has not always been consensus as to its meaning. The literal meaning of democracy comes from a combination of two Greek words, *demos* (people) and *kratos* (rule; Robertson, 1985), and at its core, “Democracy is a form of government in which the people rule” (Sørensen, 1993, p. 3). The term originated in Athens and was a part of the standard classification of “regime forms that distinguished rule by one (monarchy), several (aristocracy), and the many (democracy)” (Miller, 1987, p. 114).

However, beyond the literal meaning of democracy, there has been considerable debate over the criteria that distinguish democracies from nondemocracies. A relatively narrow definition of democracy has been offered by Joseph Schumpeter (1950), who viewed democracy as simply a method for choosing political leadership: “The democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people’s vote” (p. 260). Another, more exclusive definition is offered by David Held, who argued that “democracy entails a political community in which there is some form of *political equality* among the people” (Held, 1996, p. 1). The existence of

equal rights (and, accordingly, equal obligations) is the principal feature of political democracy.

Between the rather inclusive conception of political democracy offered by Schumpeter and the exclusive definition offered by Held is that offered by Robert Dahl (1989). For Dahl, *democracy* was an ideal type of political system in which citizens have the opportunity to (a) formulate their preferences, (b) signify their preferences to their fellow citizens and the government, and (c) have their preferences weighed equally in the conduct of government. However, since no system can fully embody democracy as an ideal type, Dahl prefers to use the term *polyarchies* to refer to existing “nonideal” democracies. Polyarchies exhibit the following characteristics:

1. Control over government decisions is constitutionally vested in elected officials.
2. Elected officials are chosen in free, fair, and frequent elections.
3. Practically all adults have the right to vote in elections.
4. Practically all adults have the right to run for elective offices.
5. Citizens have the right to express themselves freely on political matters.
6. Alternative sources of information are freely and legally available.
7. Everyone has the right to form parties, pressure groups, and other associations independent of the state. (Dahl, 1989)

One could distinguish empirically between different kinds of polyarchy in terms of two dimensions: competition for office and political participation. Systems that approach the democratic ideal (polyarchies) are characterized by high degrees of competition and high degrees of participation. Systems that have lower degrees of competition and participation are more autocratic.

However, critics of this approach argue that this conceptualization is “static” and cannot distinguish between democratic and nondemocratic regimes, but only among varying degrees of polyarchy. Furthermore, this conceptualization of democracy cannot identify how democracies emerge from nondemocratic regimes, as has occurred in most European countries. What Richard Rose, William Mishler, and Christian Haerpfer (1998) prefer is the “democracy in competition” approach to conceptualizing democracy, or the notion that democracy is defined not relative to an ideal, as is polyarchy, but relative to nondemocratic alternatives. Thus Rose et al. opt for a definition based on Juan Linz and Alfred Stepan (1996), who identify four characteristics of central importance in characterizing any regime: the rule of law, the institutions of civil society, free and fair elections, and the extent to which governors are held accountable. *Rule of law* means that no individuals, including rulers, stand above the law. *Civil society* relates to the existence of sociopolitical groups, autonomous from the state, that allow for the free articulation of popular interests and keep in check the uncontrolled growth of the state (see Diamond, 1994; Fine & Rai, 1997). *Free and fair elections* refers to the existence of real competition for office. And *accountability* refers to the extent to which those who govern are responsible to others for their political actions. None of these criteria individually is sufficient to define a democracy, or any other regime for that matter. Only in combination do these characteristics define different kinds of regimes.

Perhaps one way to combine these two very different conceptions of democracy is to think of them as measuring two very different things. On one hand, the Dahlian definition of *polyarchy* is useful in distinguishing between different varieties of democracy. Some systems are closer to the democratic ideal than others are. On the other hand, the conception of democracy favored by Rose et al. (1998) and Linz and Stepan (1996) provides the minimalist criteria for democracy and is useful in distinguishing not only democracies and nondemocracies but also regimes that are *democratizing*. From this conception, we can identify the minimal thresholds that countries pass in order to qualify as democracies—the rule of law, the development of the institutions of civil society, the existence of free and fair elections, and the extent to which those who govern are held accountable. In order to gain a comparative understanding of existing political systems, we need to account for differences between countries that have already passed the minimal thresholds that qualify them as democratic (e.g., differences in levels of participation and competition) and those countries that have only partially met the minimal

criteria for democracy, are approaching the minimal criteria, or are very far away from the minimal criteria.

If this is what we mean by democracy, what, then, do we mean by the process of democratization? *Democratization* is the process by which societies develop toward democracy. Some, like Freeman and Snidal (1982), define democratization as the extension of citizenship and the franchise. Yet this presupposes that meaningful elections take place and that political elites will abide by outcomes of such elections, which implies at least the notion that a rule of law exists and that leaders are accountable to someone. On the other hand, if we consider the minimal definition of democracy as the rule of law, the development of the institutions of civil society, the practice of free and fair elections, and the establishment of accountability of those who govern, then democratization is the process by which the rule of law, elections, and leadership accountability are established and through which civil society develops. Once established, the *expansion of democracy* involves extending the degree of competition and participation through such mechanisms as broader enfranchisement (participation) and greater competition.

To illustrate the various approaches to democracy below, this chapter deals first with some classic approaches to democracy, particularly how the “ancients” approached the concept, the idea of participatory, or direct, democracy, and then republican, or representative, democracy. It then turns to the work of more contemporary scholars, who have sought to identify the different forms that political democracy can take, especially polyarchy, majoritarian democracy, consensual democracy, consociational democracy, delegative democracy, deliberative democracy, and democratic autonomy. Last it discusses Arrow’s impossibility theorem as it relates to democracy.

Classic Approaches to Democracy

The Ancients and Democracy

The earliest conceptions of democracy are associated with the ancient Greeks. A number of factors contributed to the development of the democracy in Athens. The polis, or city-state, served as a basic unit of operation in Greece and was built on largely egalitarian values. These values were supported by three key factors.

First, the connection of lower-class citizens to the military allowed them to better their socioeconomic status, as well as to get involved in communal decision making. Second, as the Athenian polis moved toward being a “world power,” the old institutions of governance and the distribution of power were questioned. This led to a question of who should be in charge of the polis and “what role the people should play in the decisions that directly affected their safety and future” (Boedeker & Raaflaub, 1998, p. 20).

Third, the empire generated a considerable amount of income, which accumulated and allowed for extra spending

on domestic programs rather than only military. Having a steady source of income encouraged such spending, leading to “Athenians’ decision to introduce pay for juries and eventually for other political offices” (Boedeker & Raaflaub, 1998, p. 20)—an unprecedented development. However, Greek democracy was limited to freeborn male citizens with property (Thiele, 2003).

Classical scholars such as Plato and Aristotle debated the usefulness and “goodness” of pure democracy. Sometimes it was viewed as a conventional form and “sometimes as a corrupt form of popular rule in the . . . classification that included tyranny as the corrupt form of monarchy, oligarchy as the corrupt form of aristocracy, and ochlocracy as the corrupt form of government by the people” (Miller, 1987, pp. 114–115).

Plato is considered to be an opponent of democracy, even though he was a follower of the political thought of Socrates, who was believed to be a “friend of democracy and a champion of open society” (Ober, 1998, p. 156). The reason for this notion lies in Plato’s idea that democracy elevated the pursuit of freedom to the highest possible level, which ultimately leads to multiple breakdowns in the order of the society. Thus, he prefers the monarchical rule of philosopher kings (Thiele, 2003). Aristotle, a student of Plato’s, agreed with Plato that monarchy, ruled by philosopher kings, is the best possible regime. However, he realized that such a regime is impossible to maintain.

For Aristotle, democratic politics are about ruling and being ruled at the same time. Aristotle lays out three “ideal” types of regimes: monarchy, aristocracy, and polity. The corrupt counterparts of these “ideal” regimes are tyranny, oligarchy, and democracy, respectively. As was noted, monarchy was seen by Aristotle as an unstable regime that was quick to turn into tyranny, and thus, although it was the most preferred type, it was not very practical. Aristocracy is rule by the virtuous, and it turns into oligarchy if not maintained properly. Polity is the best practical regime and is a mix of aristocratic and democratic principles (Thiele, 2003).

Participatory Democracy and Direct Democracy

Although the ancients were suspicious of mob rule, the idea of participatory democracy, or direct democracy, has its roots with them. Participatory democracy, or direct democracy, can be traced back to Athens, Greece (460 BCE), where the direct rule of government was done by the people (i.e., *demokratia*) and not via an elected group of representatives (Foot, 2009). The Athenian city-state adopted this form of political system to provide its citizens an opportunity to directly participate in the state’s decision-making process. Through an assembly, citizens could directly decide and vote on “public policies that [would] govern their behavior” (Mezey, 2008, p. 1). Although the Athenian democratic form of government lacked the right of women and of slaves to vote, it still provided all adult male citizens with an increase in control over their “own

lives by allowing them to directly determine how public power [was] exercised” (Fung, 2003). As such, since there were no “representatives in the Greek system of government,” sovereignty over the laws lay primarily with male citizens, who “ruled themselves directly” (ThinkQuest Team 26466, 2009).

Currently, Switzerland and some New England town meetings closely resemble participatory democracy or direct democracy. Switzerland has 23 states, known as *cantons*, three of which are divided and known as *half cantons*. These half cantons function as full cantons by having their own constitution and legislative, executive, and judiciary branches. However, two of the half cantons perform functions that resemble the Athenian city-state political system: All adult citizens participate in the decision-making process, as in participatory democracy. The rest of the Swiss cantons use a system of representatives elected directly by citizens and who act on behalf of those citizens, constituting a republic, or a representative democracy. Most of the countries in the world today resemble a republic, or a representative democracy.

Some argue that although participatory or direct democracy allows citizens to rule themselves directly, the model may complicate and slow down the overall decision-making process (Mezey, 2008). Many critics also point out that in a direct democracy, citizens are not capable of being informed on all issues and thus are not capable of implementing various policies appropriately, and so they may instead rely on self-interest in making those decisions (Mezey, 2008). Among the many forms of democracy that have developed since the Greek and Roman civilizations, participatory or direct democracy is regarded as the type closest to the ideal form of democracy that provides citizens with full and direct participation in the decision-making process of their government (Mezey, 2008).

Republicanism, or Representative Democracy

Republicanism, or representative democracy, is also rooted in the work of the ancients. Beginning in the fifth century BCE, Romans, inspired by the Greek system of government, developed a new form of government called *republicanism* (also known as *representative democracy*) to accommodate their ever-growing population (Foot, 2009).

The difference between the Greek and Roman forms of democracy lies in the election of representatives. Specifically, in the Roman form of democracy, governmental decisions are made by an elected group of representatives (Mezey, 2008). These elected representatives “consider policy alternatives, and decide by vote” among themselves in accordance to the views of their constituents (Mezey, 2008, p. 1). In other words, in this form of government, (a) public policy is made by a representative of the citizens and not by the citizens themselves; (b) representatives are elected by citizens from groups called constituents; (c) adult citizens are able to cast a vote, and each citizen has one vote; and

(d) representatives are “accountable for their actions to those who elect them and can be replaced by next elections” (p. 2). Specifically, citizens indirectly impact political decisions by “electing and influencing the behavior of representatives who actually make public policy and control implementation” (p. 2). For example, James Madison, a founding father of the United States, and philosophers such as John Locke and Alexis de Tocqueville preferred and advocated this form of representative government, in which decisions were not made directly by citizens but were made by their elected and knowledgeable representatives. These philosophers believed that this system of representation would prevent citizens from resorting to self-interest during their decision-making process (Yarbrough, 1979).

It is important to mention that Madison aligned republicanism with representation (Yarbrough, 1979). Madison defined the republican government as one that must be democratic but not to the point that public matters must be “conducted by the citizens in person” (Yarbrough, 1979, p. 62). In general, “elected representatives would protect the right of the people better than the people themselves” (p. 62). The United States’ current system of government best resembles this form of democracy. Currently, countries such as the United Kingdom, Germany, Canada, the Netherlands, and Belgium also use this form of democracy as their governing system.

Contemporary Models of Democracy

Polyarchy

The idea of polyarchy is associated with the U.S. political scientist Robert Dahl, who was seeking an empirical way in which to measure the concept (in keeping with the introduction of the scientific method in political science as a result of the behavioral revolution in the 1950s and 1960s). In his 1963 book, *A Preface to Democratic Theory*, Dahl defined *polyarchy* as “an open, competitive and pluralistic system of minority rule” (cited in Krouse, 1982, p. 422). Dahl argued that polyarchy is “a necessary condition and foundation of democracy” (cited in Bailey & Braybrooke, 2003, p. 109). Dahl argued that “power in American politics is pluralistic,” and therefore government must account for the diversity in the population (cited in Bailey & Braybrooke, 2003, p. 103).

In his 1990 book, *After the Revolution: Authoritative Good Society*, Dahl looked at the purpose and function of polyarchy as a method of decision making (Dahl, 1990). Specifically, he emphasized that polyarchy provides greater political equality and popular sovereignty and as a democratic model best reflects participation in our modern or pluralistic society. Dahl argued that polyarchy is the basis for democracy. He also argued that pluralism is necessary, inevitable, and desirable in a polyarchy and that diversity provides individuals with more choice and leads

to self-understanding. Polyarchy is seen as the product of freedom and as generally good for human beings (Dahl, 1990). Dahl describes polyarchy as a product of democratizing nation-states and not like direct democracy as seen in ancient Athens. Polyarchy is more similar to republicanism, or the representative form of democracy.

According to Dahl, polyarchy is pivotal to the establishment of the democratic process. In his works, Dahl refers to polyarchy as a regime that must require the presence of seven political institutions in order to exist: elected officials, free and fair elections, inclusive suffrage, (an inclusive) right to run for office, freedom of expression, alternative sources of information (freedom of media), and associational autonomy (freedom of association; Bailey & Braybrooke, 2003, p. 107). In his 1972 book, titled *Polyarchy: Participation and Opposition*, Dahl emphasizes that these political institutions of polyarchy are necessary for the attainment of democracy but that ideal democracy does not exist currently in the world (Dahl, 1972). In other words, polyarchy is an imperfect, real-world substitute for full democracy (Bailey & Braybrooke, 2003, p. 107).

An important feature of polyarchy is that it promotes competition and toleration. Specifically, polyarchy accepts and tolerates a variety of views and, as such, equips citizens with an opportunity to express their opposition. Opposition parties and associations of all kinds are “good and natural” in polyarchy (Bailey & Braybrooke, 2003, p. 108). Therefore, the freedom of association and an ability of interest groups to influence governmental decision-making processes are widely accepted and encouraged in polyarchy.

Furthermore, Dahl interprets polyarchy as a system of rights. Specifically, he refers to these rights as being crucial in protecting and guaranteeing political institutions of polyarchy. Specifically, these rights include procedural rights such as (a) political equality, (b) effective participation, and (c) enlightened understanding in political and economic life (Bailey & Braybrooke, 2003).

Some scholars have argued that polyarchy is incapable of promoting democracy in societies deeply divided along cultural or ethnic lines and where “civil war is always a possibility during times of extreme conflict, especially when what is at stake is the right of a subculture to participate in governance” (Bailey & Braybrooke, 2003, p. 112).

Polyarchies not only have difficulty accommodating extreme conflict, but they may actually generate and exacerbate it. By allowing citizens to articulate their grievances freely and join associations to advance their causes, polyarchies place political weapons in the hands of people who may be culturally hostile to their fellow citizens (Bailey & Braybrooke, 2003, p. 112).

Majoritarian, or Westminster, Democracy

Majoritarian democracy is a modern form of democracy, termed the *Westminster model* by political scientist Arend Lijphart (1999) to denote the Palace of Westminster in London, where the parliament of the United Kingdom

convenes. Lijphart refers to the United Kingdom as the best example of this model.

Lijphart (1999) provides 10 distinct features to characterize this modern form of democracy:

1. Concentration of executive power in one party and bare majority: The ruling cabinet consists of a one party majority and excludes minority parties.
2. Cabinet dominance: The cabinet, composed of leaders of a cohesive majority party, can be confident of passing legislation.
3. Two party system: Government is dominated by two large parties.
4. Majoritarian and disproportional system of elections: The election system functions according to single member district plurality, or a first past the post system.
5. Interest group pluralism: Competition and conflict characterize the interest group system.
6. Unitary and centralized government: Local governments are part of the central government, their powers are not constitutionally guaranteed, and they are financially dependent on the central government.
7. Concentration of legislative power in a unicameral legislature.
8. Constitutional flexibility: For example, in the United Kingdom, there is no written constitution, and as such, Parliament can freely change policies or law by regular majorities and not by supermajorities.
9. Absence of judicial review: Since a written constitution does not exist, there is no written document by which courts can decide the constitutionality of legislation.
10. A central bank controlled by the executive: In this model, banks are controlled by the cabinet and are not independent.

Countries with this model of government tend to have homogeneous societies. This form of government can be seen in countries such as Canada, Australia, New Zealand, and most of the British former colonies in Asia, Africa, and the Caribbean after their independence (Lijphart, 1999). Although majoritarian democracy is quite prevalent in the English-speaking world, Lijphart prefers the *consensual model of democracy* in less homogeneous societies (i.e., pluralistic societies), and in fact he believes that the consensual model would be appropriate for most societies.

Consensual Democracy

In contrast to majoritarian democracy, consensual democracy is regarded by Lijphart as a better form of democracy in societies that are culturally heterogeneous (or what he calls “plural societies”). Especially in plural societies, majority rule becomes “not only undemocratic but also dangerous because minorities that are continually denied access to power will feel excluded and discriminated against and may lose their allegiance to the regime” (Lijphart, 1999, p. 32). In most deeply divided societies, such as Northern Ireland, “majority rule spells majority

dictatorship and civil strife rather than democracy” (p. 33). As such, consensual democracy is best for “less divided but still heterogeneous countries” as well as homogeneous societies (p. 33).

Lijphart (1999) provides 10 distinct features to characterize this modern form of democracy:

1. Executive power sharing in broad coalition cabinets: In this model, all or most of the important parties share executive power in broad coalition.
2. Executive legislative balance of power: There is a formal separation of power between the executive and the legislature, allowing for more independence between these two branches of government. Additionally, the legislature cannot stage a vote of no confidence.
3. Multiparty system: In a pluralist society, such as in Switzerland, parties are divided along several lines.
4. Proportional representation: This electoral system divides parliamentary seats among the parties in proportion to the votes they receive.
5. Interest group corporatism.
6. Federal and decentralized government.
7. Strong bicameralism.
8. Constitutional rigidity: A written constitution exists and can be changed only by special legislative majorities.
9. Judicial review.
10. Central bank: Some degree of independence exists for banks in monetary policy making decisions.

Consociational Democracy

Consociational democracy is a specific form of consensual democracy that Lijphart (1976) proposed in his book *The Politics of Accommodation* as a solution for societies that are deeply divided along ethnic, religious, or cultural lines. Specifically, he argued that a solution for deeply divided societies such as the Netherlands is a system of government in which groups share power within institutions. The idea of group representation is key in Lijphart’s view of achieving democracy, and the consociational model of democracy would provide for more group participation and a voice for minorities. Countries such as the Netherlands and Switzerland are the best examples of this type of democracy.

Delegative Democracy

In the 1990s, Guillermo O’Donnell (1994) introduced the idea of *delegative democracy*, which he described as follows:

[Delegative democracies] rest on the premise that whoever wins election to the presidency is thereby entitled to govern as he or she sees fit, constrained only by the hard facts of existing power relations and by a constitutionally limited term of office. The president is to be the embodiment of the nation and the main custodian and definer of its interests. (p. 60)

The policies of this democracy may not reflect the promises made by the candidate's campaign, because the candidate, once elected, is the one who decides what is appropriate for the country.

Delegative democracy occurs in formerly authoritarian states (Argentina, Brazil, Peru, Ecuador, etc.) and post-Communist countries. They meet the basic requirements of being a democratic society but are not as liberal as representative democracies "and do not seem to be on the path toward becoming" representative (O'Donnell, 1994, p. 56). These states are not consolidated or institutionalized, but they do resist regression back to authoritarianism. Delegative democracies are strongly majoritarian and hold fair and clean elections (sometimes using a runoff technique if the first round of elections does not produce a clear-cut majority). In delegative democracies, parties, the congress, and the press are normally free to voice their criticisms, unlike in authoritarian states. In certain situations, courts are able to block "unconstitutional" policies.

Delegative democracy is similar to representative democracy in that representative democracy has an element of delegation: "Through some procedure a collectivity authorizes some individuals to speak for it and eventually to commit the collectivity to what the representative decides" (O'Donnell, 1994, p. 61). However, representation requires accountability. In institutionalized societies, representatives are held accountable for their actions not only vertically, that is, to the electorate, but also horizontally, that is, to other representatives and institutions. According to O'Donnell, vertical accountability, "along with the freedom to form parties and to try to influence public opinion," is present in both delegative and representative democracies; however, the horizontal accountability attributes of representative democracy are "extremely weak or nonexistent in delegative democracies" (p. 61).

Presidents of delegative democracies make conscious efforts to disrupt the development of institutions that provide for horizontal accountability because they believe that such institutions are unnecessary impediments. Weak institutionalization in delegative democracies in turn allows the process of policy making to be swift. This increases the "likelihood of gross mistakes, of hazardous implementation, and of concentrating responsibility for the outcomes on the president," who is praised as a savior at one moment and cursed the next (O'Donnell, 1994, p. 62). In representative democracy, the decision-making process happens at a slow pace and is incremental and sometimes comes to a standstill. But the policies produced are usually less prone to gross mistakes and have a better chance of being implemented, and the responsibility for mistakes is shared among a wide range of institutions.

O'Donnell's model of delegative democracy is criticized for its failure to explain "why some presidents have been more successful than others in promoting economic reform" (Panizza, 2000, p. 738) and for undermining the importance of the politico-institutional settings in which

these reforms took place. Panizza (2000) emphasizes the importance of taking into account "the political context under which presidential power operates, the importance of coalition building and the informal and institutional constraints on that power" (p. 738).

Deliberative Democracy

Deliberative democracy is the idea that legitimate law-making stems from the public deliberation of citizens. Deliberative democracy presents "an ideal of political autonomy based on the practical reasoning of citizens" (Bohman & Rehg, 1997, p. ix). Deliberative democracy is often seen as countering *rational choice* and *liberalism* theories. Much of political action is "made up of [a] broad swath of moral conflicts" that are "not properly resolved by mere interest group bargaining" (Macedo, 1999, p. 5). Certain issues, such as affirmative action, environmental protection, or assisted suicide, cannot be resolved through rational choice argument. Many liberals too are more concerned with the fundamental rights and principles of justice than with the moral aspects of the debate. Gutmann and Thompson (1999) state that moral disagreement is ever present in politics, even under the best conditions.

Deliberative democracy promotes the legitimacy of collective decisions. Creating a feeling of legitimacy and democratic goodwill, together with fair process, creates stability in the long run. Another positive characteristic of deliberative democracy is that it encourages "public-spirited perspectives on public issues" (Macedo, 1999, p. 10). It allows the public to contemplate and think about the common good. According to Macedo's argument, deliberative democracy also promotes mutually respectful decision making, as well as the ability to correct mistakes of the past.

Despite its positive qualities, deliberative democracy has often been criticized for its idealism. Frederick Schauer believes it places too much emphasis on deliberation and "talk-based decision procedures" (Schauer, 1999, p. 18). William Simon believes that deliberative democracy's agenda is too broad and that it places too much emphasis on civility in each issue, undermining the energy of some groups that "define and constitute themselves through the assertion of their claims" (Macedo, 1999, p. 51). Last, it is criticized for presupposing a sense of closeness or solidarity among the participants, which is in fact lacking in many countries.

Democratic Autonomy

The model of democratic autonomy addresses the essential question of what it means to be democratic. Democracy as an idea offers a framework that claims that "there are fair and just ways of negotiating values and value disputes" (Held, 1996, p. 297). It is the only "grand" narrative that can legitimately "frame and delimit the competing narratives of the contemporary age" (p. 298).

Democracy does not offer a solution for all injustices and dangers. However, it does offer a first line of defense for public dialogue about general matters, as well as guiding the process of political development toward institutional paths.

The concept of autonomy implies the human ability to reason “self-consciously, to be self-reflective and to be self-determining” (Held, 1996, p. 300). This notion could not develop in a medieval worldview, because in that society, political obligations and rights were connected with property rights and tradition. As the society evolved, the concept of autonomy began to be more popular. Modern liberal society ties “goals of liberty and equality to individualist political, economic and ethical doctrines” (p. 299). These connections require the state to provide necessary conditions to allow citizens to pursue their own interests. This opinion has been largely shaped by Locke, who believed that the state exists to protect individuals’ rights and liberties and is a burden that individuals endure to secure themselves.

Democratic autonomy, therefore, requires people to enjoy equal rights and obligations “within the specification of the political framework, which generates and limits the opportunities available to them” (Held, 1996, p. 324). People should be free and equal in determining the conditions of their own lives as long as they do not impose on the rights of others. In this system, the principle of autonomy is enshrined in a constitution and bill of rights. Democratic autonomy requires open availability of information to ensure that decisions about public life are informed. It introduces new mechanisms to ensure “enlightened participation, such as voter feedback and citizen juries, increase in accountability in public and private life, and . . . [an] institutional framework receptive to experiments with organizational forms” (p. 325).

In democratic autonomy, citizens would have to accept democratic decisions in a variety of circumstances unless these circumstances violate their rights. Does that mean that democratic autonomy will require participation, making it an obligation? Held (1996) believes that this is not the case, because all citizens have the right to have a life of their own within this framework and are capable of making their own decisions to maximize their interests. The democratic autonomy model can also provide for sustained political legitimacy. The principle of autonomy can be the basis for a system that places emphasis on reducing and transforming inequalities within the society through “a double-sided process of democratization” (Held, 1996, p. 334). Only such a system may enjoy a long period of “sustained legitimation by groups other than those . . . it directly privileges” (p. 334).

Arrow’s Impossibility Theorem

Finally, any discussion of democracy would be incomplete without some reference to the work of Kenneth Arrow, and

in particular his *impossibility theorem*. This theorem states that, given several “well-known assumptions, the social orderings of particular alternatives that are meant to reflect individuals’ preferences must match the preferences of an arbitrary individual,” such as a dictator (Hansen, 2002, p. 218). According to the proof of the theorem, a social-choice rule, such as democracy, is out of the question. Dictatorship, in this case, is the only feasible form of rule.

In his proof, Arrow imagines a community attempting to make decisions about economic policy as a committee or direct democracy. The policy in question may be any other type of policy as long as it arrives at an ordering of the proposals put forward and voted on. To choose the policy, each member has to cast a vote. Each member has a preference in regard to the various proposals put forward. Arrow tries to check “the suitability of the procedure by which, from the ballots cast, the community might arrive at an ordering of the proposals which had been put forward and voted on,” and whether or not such a procedure exists (Black, 1969, p. 228).

This theorem shows that no committee procedure will be able to satisfy certain conditions that, as suggested by Arrow, this “procedure might reasonably be required to meet, and that whichever committee procedure we may choose will, for certain sets of schedules, infringe one or more of the apparently reasonable conditions” specified (Black, 1969, p. 228).

In political science, Arrow’s theorem of general impossibility found many supporters (Geanakoplos, 2005). William Riker states that voting outcomes will be different if different voting schemes for the identical set of voter preferences are tested. No voting scheme will produce a unique outcome from a given set of ordered voter preferences unless the regime is “dictatorial or manipulative” (Behrouzi, 2005, p. 64). In addition, Iain McLean stated that in a multidimensional society, the will of the people does not exist. Regardless of what option the people choose, “there’s another which a majority of the people would rather have” (Behrouzi, 2005, p. 65). These premises hint toward impossibility of direct democracy.

Conclusion

Many models of democracy appear in the contemporary political science literature. Many of these models, such as polyarchy and delegative democracy, tend to be descriptive models of reality, whereas others (such as consociationalism and deliberative democracy) tend to be proscriptive solutions to promote democracy under particular social conditions. Although each of these models assumes some basic features of democracy (such as high levels of participation, competition, and civil and political rights), they illustrate the variety of ways in which democracy can be expressed. Democracy and its conceptualization will remain important issues in 21st-century political science.

References and Further Readings

- Bailey, M., & Braybrooke, D. (2003). Robert A. Dahl's philosophy of democracy, exhibited in his essays. *Annual Review of Political Science*, 6, 99–118.
- Behrouzi, M. (2005). *Democracy as the political empowerment of the citizen: Direct deliberative e democracy*. Lanham, MD: Lexington Books.
- Black, D. (1969). On Arrow's impossibility theorem. *Journal of Law and Economics*, 12(2), 227–248.
- Boedeker, D., & Raaflaub, K. (Eds.). (1998). *Democracy, empire, and the arts in fifth century Athens*. Cambridge, MA: Harvard University Press.
- Bohman, J., & Rehg, W. (Eds.). (1997). *Deliberative democracy: Essays on reason and politics*. Cambridge: MIT Press.
- Boynton, G. R., & Kwon, W. H. (1978). An analysis of consociational democracy. *Legislative Studies Quarterly*, 3, 11–25.
- Dahl, R. (1972). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.
- Dahl, R. (1989). *On democracy*. New Haven, CT: Yale University Press.
- Dahl, R. (1990). *After the revolution: Authoritative good society*. New Haven, CT: Yale University Press.
- Diamond, L. (1994). Rethinking civil society: Toward democratic consolidation. *Journal of Democracy*, 5, 5–17.
- Fine, R., & Rai, S. R. (Eds.). (1997). *Civil society: Democratic perspectives*. London: Frank Cass.
- Foot, R. (2009). Timeline: A brief history of democracy. *Gazette*. Retrieved January 13, 2010, from <http://www.montrealgazette.com/health/Timeline+brief+history+democracy/1181297/story.html>
- Freeman, J., & Snidal, D. (1982). Diffusion, development and democratization in Western Europe. *Canadian Journal of Political Science*, 15, 299–329.
- Fung, A. (2003). Associations and democracy: Between theories, hopes, and realities. *Annual Review of Sociology*, 29, 515–539.
- Geanakoplos, J. (2005). Three brief proofs of Arrow's impossibility theorem. *Economic Theory*, 26, 1.
- Gutmann, A., & Thompson, D. (1999). Why deliberative democracy is different. *Social Philosophy & Policy*, 17, 1.
- Hansen, P. (2002). Another graphical proof of Arrow's impossibility theorem. *Journal of Economic Education*, 33(3), 217–235.
- Held, D. (1996). *Models of democracy*. Stanford, CA: Stanford University Press.
- Krouse, R. W. (1982). Polyarchy & participation: The changing democratic theory of Robert Dahl. *Polity*, 14, 441–463.
- Lijphart, A. (1976). *The politics of accommodation: Pluralism and democracy in the Netherlands*. Berkeley: University of California Press.
- Lijphart, A. (1999). *Patterns of democracy*. New Haven, CT: Yale University Press.
- Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America and post Communist Europe*. Baltimore: Johns Hopkins University Press.
- Macedo, S. (1999). *Deliberative democracy: Essays on democracy and disagreement*. New York: Oxford University Press.
- Mezey, M. (2008). *Representative democracy: Legislators and their constituents*. Lanham, MD: Rowman & Littlefield.
- Miller, D. (Ed.). (1987). *The Blackwell encyclopaedia of political thought*. Oxford, UK: Oxford Publishing Services.
- Ober, J. (1998). *Political dissent in democratic Athens: Intellectual critics of popular rule*. Princeton, NJ: Princeton University Press.
- O'Donnell, G. (1994). Delegative democracy. *Journal of Democracy*, 5(1), 55–69.
- Panizza, F. (2000). Beyond “delegative democracy”: “Old politics” and “new economics” in Latin America. *Journal of Latin American Studies*, 32(3), 737–763.
- Robertson, D. (1985). *A dictionary of modern politics*. Philadelphia: Taylor & Francis.
- Rose, R., Mishler, W., & Haerpfer, C. (1998). *Democracy and its alternatives: Understanding post Communist societies*. Baltimore: Johns Hopkins University Press.
- Schauer, F. (1999). Talking as a decision procedure. In S. Macedo (Ed.), *Deliberative democracy: Essays on democracy and disagreement* (pp. 17–27). New York: Oxford University Press.
- Schumpeter, J. (1950). *Capitalism, socialism and democracy*. New York: Harper & Row.
- Simon, W. (1999). Three limitations of deliberative democracy: Identity politics, bad faith, and indeterminacy. In S. Macedo (Ed.), *Deliberative democracy: Essays on democracy and disagreement* (pp. 49–57). New York: Oxford University Press.
- Sørensen, G. (1993). *Democracy and democratization*. Boulder, CO: Westview.
- Thiele, L. (2003). *Thinking politics: Perspectives in ancient, modern, and postmodern political theory*. New York: Chatham House.
- ThinkQuest Team 26466. (2009). *History of democracy*. Retrieved January 14, 2010, from http://library.thinkquest.org/26466/history_of_democracy.html
- Yarbrough, J. (1979). Republicanism reconsidered: Some thoughts on the foundation and preservation of the American republic. *Review of Politics*, 41, 61–95.

33

PROCESSES OF DEMOCRATIZATION

ERIC M. JEPSEN

University of South Dakota

Over the past two generations, few topics in comparative politics have generated as much research and debate as the twin subjects of *democratization* and *democratic consolidation*. Scholarship in recent decades can be seen as part of what has been a longer-standing comparative politics literature dating back to the early post–World War II period that examined subjects such as “requisites of democracy” (Lipset, 1959, p. 69) and preconditions of democratic governance (Dahl, 1971). However, it was not until what would become known as the *third wave of democratization* started building with the overthrow of the Portuguese military government in 1973 that the debates about democracy became one of the pre-eminent topics in the field. Over the course of the past 35 years, there have been a multitude of spirited debates as scholars have sought to explain what leads to democratization and how, once democratic, governments can stay democratic. What follows is an attempt to synthesize this twin body of work by examining the key arguments that have been made in the most important works within the scholarly literature.

This chapter seeks to examine the different streams of the debates in the democratization and consolidation literatures over the past decades. The first section provides a lengthy discussion of the major theoretical contributions in the study of democratization and democratic consolidation, with special attention given to the larger debates that have gripped the field over the years. The 22 key factors

thought to influence prospects for democracy are grouped into six categories and discussed at some length. The following section details the policy implications that stem from the processes of democratization and examines the broader impact these debates have had on democracy promotion and foreign policies of states, large and small. The next section sketches the likely future avenues of research that will be fruitful as further refinement of arguments about factors influencing a democratic turn (or departure) in states around the world. And the concluding section provides a summary of the bigger issues discussed in this chapter and synthesizes the wide-ranging information by listing the key conclusions that can be drawn from the literature. In addition, after the summary, there is a listing of the major works cited in this article and other scholarship that would be of interest for a more in-depth examination of the topic of democracy and democratic consolidation in the field of comparative politics.

Theory, Application, and Evidence

Theoretical debates about democratization and democratic consolidation have been some of the most contested and plentiful of any of the theoretical debates in comparative politics and political science. Over the course of the post–World War II period, scholars have researched and theorized as to the factors that best explain the turn toward

democracy in nondemocratic regimes and the factors that lead to the continuity of democracy and the point at which a state can be considered “safely democratic” or a consolidated democratic country. This section of the chapter outlines the 22 explanations that are thought to be the most important ones influencing democratization and the movement toward democracy around the world. These factors are those that are thought to have influenced “early” democratizing countries like the United States and France, as well as impacting those cases that could be considered to be part of the grouping of countries that moved toward democratic rule in the latter part of the 20th century.

Samuel Huntington (1991), in his seminal book on democracy and democratization, has called these historical periods in which large numbers of countries around the world democratize the three “waves” of democratization. The first wave is considered to be the “long wave,” which took place from the late 18th century until the aftermath of World War I. Although the states that took the democratic path then were few, it was during this period that what would become key long-standing democratic states in North America, Europe, and other parts of the English-speaking world democratized.¹ The second wave is considered to be the time that began in the aftermath of World War II and continued for the better part of two decades.² The second wave was largely the result of decolonization and the systemic effects resulting from allied occupation of and involvement in certain countries and territories in the aftermath of large-scale war in Asia and Europe in the 1930s and 1940s. According to Huntington’s arguments, the third wave of democracy started in southern Europe in 1974 and continued into the 1990s.³ This period is associated with the ending of authoritarian regimes throughout almost all regions of the world and especially the fall of the Communist regimes in Europe in the latter part of the 1980s and early 1990s.

This chapter breaks the explanatory factors thought to lead to (or promote) democratization into six different categorical groups. Recent decades have seen comparative scholars argue that a multitude of factors have led to countries’ becoming democratic. Likewise, many arguments have attempted to account for the consolidation of democracy in certain cases but not in others. Although there have been bitter disagreements in the subfield as to which variables are the most important, this chapter attempts to steer clear of “taking sides” and instead provides a template of the major debates within the literature. The variables discussed in this chapter are based on discussions and synthesis by Huntington (1991), Linz and Stepan (1996), Sörensen (2008), and Tilly (2007) and represent the past several decades of scholarship.

The first category of factors that are thought to create a positive environment for democracy and democratization is *modernization*. This is the classic explanation for democracy and dominated our understanding in the early decades of the debate. The second category includes those variables that can be grouped into *economic preconditions*.

These are factors that have distinct causal mechanisms related to economics or economic development. The third category contains all explanations that are based on what may be labeled *social preconditions*. This category includes a range of explanatory factors, from social structure to metalevel cultural explanations. The fourth group of factors is factors dependent on *timing, sequencing, and politics*. This category includes arguments that have some distinct and explicit treatment of the temporal element of democratization throughout the world. The fifth category includes *agency and advocacy rooted explanations*. These are factors rooted in the activities of social or civil organizations. The sixth category covers explanations related to *external and foreign effects* and *democratic experiential arguments*. All these explanations consider the impact of nondomestic sources of influences on democracy and democratic consolidation.

Modernization

One of the longest-standing categories of explanations as to why countries democratize has to do with what can be called modernization or modernization theory. While this has been one of the most contested and tested arguments in the literature, even after 50 years it remains an important area of research into the sometimes necessary (but not necessarily sufficient) conditions for democracy and democratic governance in states. First formulated in the 1950s, during a period that much later would come to be thought of as the end of the global second wave of democratization, modernization theory suggested that certain factors were necessary for democracy to take root in states across the world. Basically, Lipset (1959), and later others, argued that certain factors such as *high levels of wealth, higher degrees of urbanization, better educated populations, and increased industrialization* would lead to pressures for democracy and democratization. All these factors that were typically associated with what was labeled *modernization* were thought to be the crucial underpinnings of the movement toward democracy in much of the world. As the state became more capable, thanks to economic development and the rationalization of state–society interactions due to modernization, democracy would become more viable and increasingly effective. These supporting pressures were thought to allow for “democratic awakenings” (and later consolidation) in states around the world.

However, Lipset and the early modernization theorists were not without their critics. Scholars in later years would argue that modernization theory was reliant on a handful of cases in North America and Europe and that the processes Lipset and others described did not necessarily always play a supportive role. Scholars would come to criticize many of the suppositions and causal linkages of modernization theory. Some would reverse the causation and say that in certain contexts, authoritarian regimes could be supported or empowered by the very factors thought to sustain democracy under the modernization model. For example,

Rueschemeyer, Stephens, and Stephens (1992) would make the case that this growing state strength could allow authoritarian regimes to push back against democratic forces within a country. O'Donnell (1973) reversed the directional arrows altogether and made the case that the modernizers had it all wrong and that under certain conditions, a strengthened state and antidemocratic elites could forge an alliance to overturn democracy precisely because of modernization. He argued that the nature of lopsided economic development and empowerment of certain business and administrative factions in Latin American countries meant that an increase in incomes and gross domestic product could actually undermine democracy. Later, other scholars would argue that modernization arguments that democratic advocates would be activated in the middle classes as the result of improving social and economic conditions were faulty as well. Critics have also claimed that under certain cases of modernization, there could be an increase in societal tensions along group lines (Diamond, 1989) or an increase in antidemocratic tendencies among elites in modernizing societies (Moore, 1966).

Economic Preconditions

Closely related to modernization theory, economic factors have been thought to impact the likelihood of democratization in three particular ways. It is helpful to look more closely at each of them in isolation and examine the ways that previous comparative political theorizing has posited that economic factors are important in the movement toward democracy and democratic consolidation. The first economic variable that scholars have thought significant in democratization is *wealth*. This is typically considered to be measured by gross domestic product per capita so as to account for differential population size across nations. Lipset (1959) was among the first scholars to explain the manifest prevalence of wealthier countries among democratic countries. In later years, the thesis that wealth leads to democracy would undergo some amount of revision as scholars began to test the linkages put forward by Lipset and the modernization theorists.

As Przeworski and Limongi (1997) and Przeworski, Alvarez, Cheibub, and Limongi (2000) have detailed, we must break down our discussion of wealth and its impact into different temporal relationships with the likelihood of democracy. For example, they argue that we should think of the relationship as being either *endogenous* or *exogenous* when we think of the mechanisms by which wealth is related to prospects for democracy. Their argument is that there is little support for the theory that higher amounts of endogenous wealth lead to greater likelihood of transition. Yet there does seem to be support for an exogenous relationship—that once democratic, countries that are wealthier tend to stay democratic to a greater degree than do those countries that are less wealthy. Although the distinction might seem minor, it gets at the heart of the debate about which processes and factors lead to democratization

and to democratic consolidation. Many scholars argue that the effects are seemingly minor on the democratization side but paramount on the consolidation side. Later scholars have questioned whether these findings hold up for different regions and in different times. In certain regions in certain times, there is more likely to be a transition from authoritarian rule to democratic rule in any given year and in any given case as a country becomes wealthier.

Another notable exception to the wealth-causes-democracy line of inquiry in the literature is the antidemocratic tendencies seen to result from “bad” types of wealth. Specifically, an area of research suggests that countries with strong dependence on a resource that generates great wealth, such as oil, have greater difficulties transitioning to democracy and sustaining democracy over the longer term. Scholars in recent decades have used this argument to make the case that the problems with democracy over the years in countries as divergent as Saudi Arabia, Venezuela, Russia, Mexico, and Nigeria are the result of oil wealth’s “cursing” democratic governance. In a related area of research, there is a long-standing finding in the economic literature that natural resource abundance can, under certain conditions, lead to depressed development and uncompetitive industry because of economic distortions caused by the very reliance on a resource that is generating a large amount of foreign currency and goods. The basic argument in terms of democracy is that the “easy wealth” captured by the state from the export of the natural resource (oftentimes thought to be most pernicious in the case of oil) leads to antidemocratic tendencies in the state because of increased reliance on corporatist and clientistic governing mechanisms. Although there are cases where this problem is thought to be minimal (Norway), many scholars make a clear distinction between “good wealth” and “bad wealth.”

The second economic precondition thought to condition the likelihood of democratization and prospects for democratic consolidation is the effects of *capitalism and free market economies*. Scholars such as Almond (1991), Dahl (1989), and others have argued that key supporting features of democracy built into the practice of capitalism and the functioning of free and open markets serve to reinforce democracy and democratic practices. Almond, in his seminal piece on capitalism and democracy, reviewed the various arguments that have sought to link the practice of capitalism with support for democracy.⁴ As well, Dahl has argued that the linkage is clear and that definite patterns of support result from economic development that comes about in capitalist systems. In addition, capitalism also allows for the creation of other independent factors thought to support the turn toward democracy. Specifically, it is thought to result in many instances in an expansion of the middle class and empowerment of social and economic interests via capitalist development. These are then thought to aid in democratization and democratic consolidation.

The third economic precondition thought to positively influence chances for democratization and democratic consolidation is the *equitable distribution of, and access*

to, land and resources. The roots of this argument are based on the expectation that more equitable and egalitarian societies provide the support necessary for democracy to take hold. As Moore (1966) and Dahl (1989) have suggested, societies with less polarized resource distribution and less concentrated landholding patterns are thought to reinforce elite and interest group participation in stabilized “rules of the game” that democracy provides. In more polarized, less egalitarian land and resource situations, it is thought that a winner-take-all circumstance could lead to destabilizing and antidemocratic tendencies. Although both Moore and Dahl acknowledge that equitable distribution of land and resources is not a sufficient cause of democracy and democratic stability, it is thought that already polarized group-based societies could dangerously reinforce conflict that could harm progress toward democratization. Scholars in recent years have pointed to the lack of this condition in many cases around the world as one of the key reasons for the few “new” cases of democracy in the past decade or so.

Social Preconditions

The next category of causal mechanisms that have been thought to be supportive of democratization and democratic consolidation contains those that can be labeled social preconditions. The first of these is thought to be that the *absence of a history of feudalism* (as well as non-European feudalistic traditions) is better for the flourishing of democracy. Moore (1966) counts the longer-term effects of feudal social structures as one of the more important factors lessening the prospects for democracy in certain cases, with the basic tenets of the argument suggesting that a rigid social system casts a shadow and has debilitating effects in the longer term. While this condition is associated with some of the others discussed in this chapter (especially land use patterns and the degree of pluralism within society), it is distinct in that scholars have explored the narrower effects of previous feudal relationships on later relative strength of antidemocratic forces in certain cases.

The second social precondition is what can be labeled a “strong bourgeoisie,” in the words of Moore (1966), or a *strong middle class*. Although functionally different in classical definition, a strong middle class and a strong bourgeoisie have come to be discussed in unison in recent years. These concepts are also related to other social conditions discussed in this section. Moore’s comparative history of democratization in several important cases led him to argue that one of the most crucial inputs for successful democratization (and longer-term democratic sustainability) was the existence of a linchpin supporter of democracy in the form of a bourgeoisie class. The bourgeoisie serves as the stabilizing influence over and in favor of regularized and transparent political processes that can be best handled by democratic forms of government. After democratization, the bourgeoisie serves as the primary sustainer of consolidation in the face of antidemocratic tendencies

among elites and lower classes, both rural and urban. Similarly, as argued by Putnam (1993) in his analysis of relative democracy across the regions of Italy, the middle classes can serve as a buffer between the extremes of the political landscape and lead to a necessary underpinning of democracy and democratic governance.

The third social precondition that is argued to be important to democratization and democratic consolidation is *Protestantism*. The root of this argument can be found in the much earlier arguments of the early-20th-century German sociologist Max Weber and his arguments about the supportive framework that Protestantism provides for capitalism. In later scholarship, researchers such as Wiarda (1982) argued that Latin America’s difficult experience with democracy and democratic stability during the post–World War II period was the result of the detrimental effects of Catholicism on the social structure in the region. Many of the arguments that have been made about Protestantism serving as a source of support for democracy are nested in reasoning that traces how it leads to higher levels of accountability and citizen participation in political life, which enables democracy to take hold in different cases. Nevertheless, these arguments are much contested by many scholars, and the longer-term viability of some of the findings has been limited.

The fourth social precondition that scholars have cited as being instrumental in supporting democratization and democratic consolidation is the basket of arguments that concern *political culture, civic culture, and values*. This line of inquiry argues that there are certain societies that have, at the macrocultural level, attributes and tendencies that support the democratic project. Almond and Verba (1963) and Inglehart (1990) make arguments that this is a key explanatory variable that determines the relative likelihood of successful democratization and consolidation. Scholars that have made these types of cultural arguments contend that these core, slow-moving, society-wide values greatly impact a country’s experience with democratic governance. In the end, it is thought by proponents of this line of thinking that unsupportive values and political cultures will lead to repeated democratic setbacks and reversals, with prospects for long-term democracy possible only after a change in those values and the macroculture.

The fifth and final social precondition variable thought to impact prospects for democracy is that *cultural homogeneity* is a positive force for democracy and democratic consolidation. This argument suggests that in situations in which there are significant societal cleavages or ingroup–outgroup patterns of polarization, democracy is difficult to promote and sustain. Scholars point to the early outposts of (relatively) homogeneous countries in North America and Europe as transitioning to democracy more quickly, and trace out the causes of democratic reversals in Asia and Latin America as being the result of group conflict and heterogeneous societies. There have been glaring exceptions to this precondition, with highly

diverse countries such as India having successful democratic track records in the latter half of the 20th century.

Timing, Sequencing, and Politics

An additional category of explanations for successful democratization and democratic consolidations focuses on distinctly political patterns, sequencing, and timing. The first argument is that the *development of contested politics before the expansion of suffrage and participation* leads to a greater chance of a successful democratic experience. The logic here is that newly democratizing countries cannot handle the system shock of having the political system open up all at once and that a lasting commitment to democratic practices can result only after years or decades of successful political competition among a narrow band of a polity. Scholars point to the United States and Great Britain and their respective many-decades-long opening up of political participation by group and gender well after regularized political competition among a smaller band of society (in each case by landed, white males at the start). Although this argument is difficult to advocate in terms of a policy prescription, researchers have pointed to it as a source of democratic breakdown in Latin American and Asian cases in particular.

A second trait in this category is the importance of *low levels of civil violence, polarization, and extremism* in the likelihood of democracy and consolidation. The logic here is fairly straightforward: In conditions of instability and violence, actors at the elite and mass levels are less likely to commit to regularized and open democratic governance when there are immediate threats to their lives and livelihoods. Because democracy and democratic consolidation require commitments to future payoffs despite possible near-term loss of political power, violence and extreme polarization can stop a turn toward democracy in its infant stages. Until there is a brokered and longer-term cessation of these types of problems, scholars have argued, democracy is unlikely to be successful.

A third trait in this category is the *fairly uncomplicated experience* argument. Theorizing and empirical research have consistently demonstrated that democratization is rarely successful on the first go-round (Przeworski et al., 2000). Countries that have had prior experience with democracy are more likely to revert to it sometime in the future than are countries with no sort of democratic tradition or “practice.” Two sets of logic in the literature support this trait. One, if a country has had some form of democratic experience, democracy will be perceived by groups within that society as a possible “real” option in the future. And two, previous iterations of democracy have likely left behind building blocks of future democracy: political parties and civic groups that likely have a permanent openness toward democratic governance.

A fourth trait in this category is Tilly’s (2007) recent conception of *trust networks and categorical equality*.

Specifically, he has argued that the subjugation of the state to public politics, along with expanding popular control over that political game, results in regularized control over governance. This regularized control over governance leads to the formation of trust networks as people become willing to abide by the set rules of the democratic game. And, in a supporting role from the other direction, the state monitors antidemocratic tendencies and groups and seeks to eliminate these threats before they undermine the democratic project.

Agency and Advocacy

The next category of explanations relates to specific actions and work by elements (individuals and groups) within society. It is thought that agency and specific activities can play a supporting role in democratization and democratic consolidation. The first of these agency-based explanations is what has come to be known as the “*elite pacts*” argument. First formulated by O’Donnell and Schmitter (1986), this argument holds that in order for transitions to democracy to be successful and long lasting, top-society groups must forge agreements about democracy’s ability to regularize political competition among competing elite groups. These privileged groups, therefore, would agree not to “kick over the table” (i.e., support nondemocratic and authoritarian governments). The argument here is that “pacted democratization” is necessary, especially in societies and cases in which elite groups have had extreme and polarized patterns of political interactions. O’Donnell and Schmitter pointed to many Latin American and southern European cases as having democratized successfully, primarily because of top-level agreements among politically powerful groups.

The second agency-based explanation for successful democracy and democratic consolidation revolves around arguments that *active civil societies* create the necessary conditions for the turn toward democracy and, later, its consolidation. This line of inquiry and theorizing is rooted in the type of analysis provided by Alexis de Tocqueville, the 19th-century French writer and political observer who identified the richness and diversity of civic life in the United States as the critical underpinning of its democratic life. The thrust of the argument is that an activated and engaged mass public will deliver sustained democratization pressures via the measures resulting from regularized, collective interaction.

The third explanation in this category is related to the richness of civil society argument noted above. However, this explanation is rooted in the impact of the *character of the social groups* themselves. The argument is that all social and civic groups are not created equal. Some have greater impact on prospects for democracy and democratic consolidation. Mass democracy becomes more likely when microlevel democracy is present in groups that have a direct or instrumental role in political life. Political parties and policy-oriented social groups that

operate in a democratic manner are likely to lead to “spillover democracy” in general political life.

External Actors

The final set of explanations for democratization’s birth and growth all revolve around the role and legacy of external actors. First, it has been argued that *intervention by a foreign party*, under certain conditions, can lead to democratization. Likewise, according to some scholars, support and prodemocratic pressure from abroad can lead to increased likelihood of democracy’s taking hold. The successful and rapid democratization of Japan and West Germany following World War II is pointed to as a successful example of foreign influence on newly democratizing countries. Although the involvement of a foreign power in a country’s political affairs is oftentimes antidemocratic, there is some evidence that clearly prodemocratic influence from an outside force can lead to increased chances for democratic survival.

A second type of external argument is called the *diffusion effect*. Specifically, this argument suggests that the diffusion of democracy by geographical or temporal clusters has its own stand-alone effect separate from the component parts or similarities of cases. In other words, a distinct “wave effect” accounts for part of the democratization that occurs in similar time periods and in certain regions of the world. Whether this effect is due to creation of possible options in the minds of key actors within countries or reflects a metalevel supporting environment, many scholars have suggested that diffusion helps explain why and when democratization occurs and consolidates.

A third type of external argument is that something about the *colonial experiences* of certain states conditions pathways to democratic governance. In particular, scholars have argued that the British colonial experience is distinct from other colonial experiences (namely, Spanish, Portuguese, French, etc.) and that former British colonies are more likely to be democracies than are other former colonies. The key argument here is that the specific components of the British colonial experience—especially professionalized bureaucracies, elements of the Westminster parliamentary system, and elite education systems—cast a long shadow that increases the likelihood of successful democratization and democratic consolidation.

Policy Implications

A host of policy implications result from the wider theoretical literature on democratization and democratic consolidation. Lessons from the research could be applied by individual countries, foreign powers, and nongovernmental organizations. Although the democratization literature contains some contradictory arguments, by and large it identifies a baseline level of standard supporting components for increasing the numbers of democracies in the

world. Trends in many of these areas can be found in the yearly *Freedom in the World Report* by Freedom House (www.freedomhouse.org). This section discusses briefly the key variables that should be improved in order to increase the number of democracies in the world and to stabilize and consolidate those already in existence.

On the economic front, it is generally thought that economic development and the alleviation of poverty will lead to increased prospects for democracy. Incorporation of some elements of the market economy and modernization will also lead to a supportive environment. In some cases, there should be a certain amount of land reform or redistribution if the society has skewed distributions of wealth, and natural resources should be held in a way that benefits the vast majority of the population. None of these policy suggestions are sufficient for democracy and the consolidation of democracy, but the theoretical literature generally agrees that these types of reforms and policies would be supportive.

On the social and political fronts, it is thought that all efforts should be made toward alleviating the tensions associated with civil strife that is the result of group difference. Institutional reforms should be adopted that would allow for less friction in a country’s political life. As well, it is believed that civic groups and nongovernmental organizations should be empowered so as to provide a further building block of regularized and peaceful political organization and interaction.

Future Directions

Research in comparative politics and international relations on democratization and democratic consolidation has a long and rich history of theoretically informed scholarship. In addition, the subfield faces challenges and associated research questions. First, future research needs to do a better job of explaining the conditions under which these 22 factors positively reinforce democracy and democratic consolidation. Does their effect vary to a large degree by region and time? Research must also address the way these processes relate to each other and what the policy implications are for governments. Second, future research needs to explain democratic reversals and backsliding. The slowing of democratization around the world in the past decade has also corresponded to instances of democratic death and decline in certain cases. What can existing theory say about the reasons for these problems and what are the warning signs of a “failure to consolidate”?

As we move toward the 40th anniversary of the start of the third wave, we seem to have a limited understanding of what constitutes full consolidation, instead of what might merely be “not failed as of yet.” Tilly (2007) has attempted to sketch out an agenda for what research into “de-democratization” might look like, but there is a long way to go before we have a comprehensive framework by which we can explain what leads to the formative stages of problems

in democracy. And finally, it seems that with the “stabilizing” (or lack of democratic gains) of the number of democratic states around the world in recent years, we do not have a solid understanding of the most important factors in determining where democracy’s next footholds will likely be and when they might take place. Although comparative politics and students of democratization hesitate to engage in predictive and proscriptive analysis, because of the real security issues now faced around the world, a proper engagement of this line of inquiry must occur as theory building continues.

Conclusion

This chapter has surveyed the divergent strands of 50 years of comparative politics scholarship on the topic of democratization and democratic consolidation and concentrated on the most influential and most debated factors over recent decades. These factors coalesce into six categories: modernization; economic preconditions; social preconditions; timing, sequencing, and politics; agency and advocacy; and external actors and experience. Identifying these categories highlights the causal mechanisms that scholars think have increased the numbers of democratic states around the world. In addition, many of these factors promote the consolidation of democracy once it has been established in a particular case. Although many of the factors have had their respective advocates and detractors over the years, nearly all of them continue to be thought important, both in the academic world and in the real world of governance.

We can also see the broad outlines of what areas of research and theorizing still need attention after the past half century of debate. Lipset published “Some Social Requisites of Democracy” in 1959, and since that time, comparative politics research has made great strides in explaining the relationship between various factors and democratic governance around the world. Because of geopolitical realities and the continued debates surrounding democracy promotion and economic development around the world, it is likely that an examination of the processes of democratization and democratic consolidation will continue to be a key topic in comparative politics. Large numbers of scholars continue to ask important questions, and the nature of politics in many states around the world suggests that the real-world implications will continue to drive the debates forward. In all likelihood, the next half century will see further important contributions to this line of inquiry, and there will surely be additional sophistication in the models seeking to identify the causal mechanisms of democratization and democratic consolidation.

Notes

1. Countries that Huntington lists as democratizing during the first wave are Australia, Canada, Finland, Iceland, Ireland,

New Zealand, Sweden, Switzerland, the United Kingdom, the United States, and Chile. Some other countries democratized during the first wave but later reverted to nondemocratic rule (notably Austria, Belgium, Colombia, Denmark, France, Germany, Italy, Japan, the Netherlands, and Norway).

2. Countries in the second wave of democratization are Botswana, Costa Rica, Gambia, Israel, Jamaica, Malaysia, Malta, Sri Lanka, Trinidad and Tobago, and Venezuela. As with the first wave, several countries democratized during this period but later reverted to nondemocratic rule (notably Bolivia, Brazil, Ecuador, India, South Korea, Pakistan, Peru, the Philippines, and Turkey).

3. He argues that the third wave started with the fall of military regimes in Portugal and Greece and later spread to many parts of the world (notably Bulgaria, the Dominican Republic, El Salvador, Guatemala, Honduras, Mongolia, Namibia, Nicaragua, Panama, New Guinea, Romania, Poland, and Senegal).

4. To be fair, however, Almond (1991) goes on to explore the literature that suggests that capitalism can be detrimental to the development of democracy. He also explores the arguments associated with the causal arrows pointing in the opposite direction: democracy supporting capitalism and democracy undermining capitalism.

References and Further Readings

- Almond, G. A. (1991). Capitalism and democracy. *PS: Political Science and Politics*, 24, 467–474.
- Almond, G. A., & Verba, S. (1963). *The civic culture*. Princeton, NJ: Princeton University Press.
- Baldez, L. (2003). Women’s movements and democratic transition in Chile, Brazil, East Germany, and Poland. *Comparative Politics*, 35(3), 253–272.
- Bermeo, N. (1994). Sacrifice, sequence, and strength in successful dual transitions: Lessons from Spain. *Journal of Politics*, 56, 601–627.
- Bermeo, N. (2003). *Ordinary people in extraordinary times: The citizenry and the breakdown of democracy*. Princeton, NJ: Princeton University Press.
- Bernhard, M. (2005). *Institutions and the fate of democracy*. Pittsburgh, PA: University of Pittsburgh Press.
- Bohman, J., & Rehg, W. (Eds.). (1997). *Deliberative democracy: Essays on reason and politics*. Cambridge: MIT Press.
- Boix, C., & Stokes, S. C. (2003). Endogenous democratization. *World Politics*, 55, 517–549.
- Bunce, V. (2000). Comparative democratization: Big and bounded generalizations. *Comparative Political Studies*, 33, 703–734.
- Bunce, V. (2003). Rethinking recent democratization: Lessons from the postcommunist experience. *World Politics*, 55(2), 167–192.
- Bunce, V., & Wolchik, S. (2006). *Favorable conditions and international support: The wave electoral revolutions in post communist countries*. Unpublished manuscript.
- Collier, D., & Levitsky, S. (1997). Democracy with adjectives: Conceptual innovation in comparative research. *World Politics*, 49, 430–451.

- Dahl, R. A. (1971). *Polyarchy*. New Haven, CT: Yale University Press.
- Dahl, R. A. (1989). *Democracy and its critics*. New Haven, CT: Yale University Press.
- Dahl, R. A. (2000). *On democracy*. New Haven, CT: Yale University Press.
- Di Palma, G. (1990). *To craft democracies: An essay on democratic transitions*. Berkeley: University of California Press.
- Diamond, L. (1989). Introduction: Persistence, erosion, break down and renewal. In L. Diamond, J. J. Linz, & S. M. Lipset (Eds.), *Democracy in developing countries: Asia*. Boulder, CO: Lynne Rienner.
- Ekiert, G. (1991). Democratization processes in East Central Europe: A theoretical reconsideration. *British Journal of Political Science*, 21, 285-313.
- Epstein, D. L., Bates, R., Goldstone, J., Kristensen, I., & O'Halloran, S. (2006). Democratic transitions. *American Journal of Political Science*, 50(3), 551-569.
- Geddes, B. (1999). What do we know about democratization after twenty years? *Annual Review of Political Science*, 2, 115-144.
- Gutmann, A., & Thompson, D. (1998). *Democracy and disagreement*. Cambridge, MA: Belknap Press.
- Gutmann, A., & Thompson, D. (2004). *Why deliberative democracy?* Princeton, NJ: Princeton University Press.
- Huntington, S. P. (1991). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Inglehart, R. (1990). *Culture shift in advanced industrial society*. Princeton, NJ: Princeton University Press.
- Keane, J. (1988). *Democracy and civil society*. London: Verso.
- Levitsky, S., & Way, L. A. (2002). Elections without democracy: The rise of competitive authoritarianism. *Journal of Democracy*, 13(2), 51-65.
- Lijphart, A. (1977). *Democracy in plural societies: A comparative exploration*. New Haven, CT: Yale University Press.
- Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty six countries*. New Haven, CT: Yale University Press.
- Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and post Communist Europe*. Baltimore: Johns Hopkins University Press.
- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53, 69-105.
- Lipset, S. M. (1960). *Political man: The social bases of politics* (Expanded ed.). Baltimore: Johns Hopkins University Press.
- Moore, B. (1966). *Social origins of dictatorship and democracy: Lord and peasant in the making of the modern world*. Boston: Beacon Press.
- Munck, G. L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, 35, 5-34.
- O'Donnell, G., & Schmitter, P. C. (1986). *Transitions from authoritarian rule: Tentative conclusions about uncertain democracies*. Baltimore: Johns Hopkins University Press.
- O'Donnell, G. A. (1973). *Modernization and bureaucratic authoritarianism*. Berkeley: University of California Press.
- Olson, M. (2000). *Power and prosperity: Outgrowing communist and capitalist dictatorship*. New York: Basic Books.
- Perez Diaz, V. M. (1993). *The rebirth of civil society: The emergence of democratic Spain*. Cambridge, MA: Harvard University Press.
- Pharr, S. J., & Putnam, R. D. (2000). *Disaffected democracies: What's troubling the trilateral countries?* Princeton, NJ: Princeton University Press.
- Przeworski, A. (1991). *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. New York: Cambridge University Press.
- Przeworski, A. (1995). *Sustainable democracy*. New York: Cambridge University Press.
- Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development: Political institutions and well being in the world, 1950-1990*. New York: Cambridge University Press.
- Przeworski, A., & Limongi, F. (1997). Modernization: Theories and facts. *World Politics*, 49, 155-183.
- Putnam, R. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Rueschemeyer, D., Stephens, E. H., & Stephens, J. D. (1992). *Capitalist development and democracy*. Chicago: University of Chicago Press.
- Sartori, G. (1987). *The theory of democracy revisited*. Chatham, NJ: Chatham House.
- Sartori, G. (1997). *Comparative constitutional engineering: An inquiry into structures, incentives and outcomes*. New York: New York University Press.
- Schmitter, P. C., with Karl, T. L. (1994). The conceptual travels of transitologists and consolidologists: How far east should they attempt to go? *Slavic Review*, 53, 173-185.
- Schumpeter, J. A. (1950). *Capitalism, socialism, and democracy* (3rd ed.). New York: Harper & Row.
- Sörensen, G. (2008). *Democracy and democratization: Processes and prospects in a changing world* (3rd ed.). Boulder, CO: Westview Press.
- Tilly, C. (2007). *Democracy*. New York: Cambridge University Press.
- Wiarda, H. (1982). *Politics and social change in Latin America: The distinct tradition* (Rev. 2nd ed.). Amherst: University of Massachusetts Press.

34

COMPARATIVE METHODS

MICHELLE C. FLORES

University of Southern California

As a subdiscipline of political science, comparative politics aims to explain and understand the dynamics of political power as practiced throughout the world. In pursuit of this goal, comparativists have developed a range of methods to compare the large number of vastly different political systems they study. While philosophers, historians, and theologians have long crafted political theory in a systematic fashion, the establishment of modern political science departments and the rapid increase in their number during the 20th century inspired a fruitful debate about the appropriate means to carry out comparative political research. In the early 21st century, there is growing recognition of the necessity of multiple methods, and recent methodological debates have centered on the best ways to enhance dialogue between scholars from different methodological backgrounds who nevertheless share substantive concerns.

This chapter provides an overview of comparative methods as understood by their practitioners. It presents a number of alternative approaches, discusses their implications, and shows how these approaches have been used in exemplary works in the field. The chapter ends with a discussion of current trends in comparative methodology and how they might impact the future of the discipline.

Comparative Methods: Qualitative, Quantitative, and Formal Methodologies in Comparative Research

Taking the natural sciences as its model, political science has sought to create theories to explain and predict various aspects of political life. Indeed, political scientists have striven to shape their craft scientifically by putting in place and advocating systematic research processes aimed at cumulating knowledge. In this sense, the choice of method is but one step in a larger research process that usually includes a clear delineation of the research question, an examination of the existent theory related to the problem, a description of the data to be used, a method of data analysis, and discussion of the potential contribution to theory. The sum of these parts is referred to as the research design, and comparativists generally agree that it should be both logically consistent and justified by the problem it studies. Therefore, in assessing the range of comparative methods, it is important to look at how comparative methods fit with various aspects of research design.

The most influential early work on research design was Adam Przeworski and Henry Teune's (1970) *Logic of Comparative Social Inquiry*. Their work aimed at designing research that would develop general social theory by confirming, through comparative research, hypothetical

statements that replaced proper names of social systems with names of variables. They posited a basic distinction between what they called *most similar* and *most different systems* research designs. In most similar systems research designs, cases are chosen on the basis of assumed similarities at the systemic level (state, culture, nation, etc.), whereas in most different systems designs, the type of cases and the level of analysis emerge from the analysis of theoretically relevant factors in data that assume the homogeneity of all units. Although Przeworski and Teune did not deny that there was some value in most similar systems designs, their delineation of comparative research was particularly rigid insofar as it asserted that the most different systems design, the definition of whose units was based on a random multistep sample of all social systems, was the *only* research design that could allow universal generalizations. Nevertheless, their argument was enormously influential and sparked an invaluable debate within the field about the goals of research and the importance of research design.

Przeworski and Teune's argument went well beyond the matter of choosing cases, however, and sought to emphasize the scientific quality of the comparative method. Arend Lijphart (1971) furthered this logic, depicting the comparative method as a way of achieving scientific explanation, albeit one with certain limitations. Chief among the difficulties facing comparativists, Lijphart contended, was constructing parsimonious theories based on research that inherently involved many variables but few cases, especially cross-national research. This difficulty was not seen as debilitating, however, and many of the ways that Lijphart suggested to mitigate the problem—including conceptual and statistical techniques for reducing the number of variables and increasing the number of cases—continue to be used today (see section titled "Scope" below).

More recent methodological debates, however, center less on justifying a scientific approach to political phenomena than on arguing a best fit between research question and the types of data that will be gathered, how they will be analyzed, and the relationship between data analysis and theory. Although the mainstream methodology literature in comparative politics continues to advocate a quantitative, statistical approach to studying comparative politics, there is growing recognition that the methodological landscape has become far more complex. It can be roughly divided into two categories: *empirical* and *formal* methods. Empirical methodologies are largely divided between *quantitative* and *qualitative* traditions, and the formal methods used in comparative politics are dominated by *game theoretic models of rational choice theory* (Laitin, 2002).

Research as a Mediated Encounter Between Theory and Fact

Whether perceived as a constant dialogue or as one controlled instance, comparative political research can be

usefully described as the researcher's fruitful encounter with theory and fact. Comparative methods mediate this encounter, providing researchers with systematic ways to produce knowledge based on what was previously understood about an issue and what can be observed in the world. They help the researcher explain connections, concepts, and causes that are not observable without systematic analysis. Thus, comparative methods are at the center of the systematic processes political scientists use to facilitate the creation and transmission of knowledge.

The choice of method impacts or is impacted by the decisions scholars make at every point in the research process, from choosing the research question to presenting their conclusions. Of course, there is a great deal of variation within methodological traditions, as well as some overlap in their application and potentialities. In fact, the differences presented here are not rigid, and much of the methodological innovation in the field rests on the ability of researchers to create internally consistent research designs that cannot be neatly categorized on either side of traditional methodological divisions. Nevertheless, for the purpose of this chapter, it is useful to sketch these ideal types based on their use in the discipline. What follows is a consideration of the role of comparative methods as a mediator in three aspects of research: theory generation and the goals of research, methods of analysis, and theory assessment.

Methods of Theory Generation and Goals of Research

Theory generation in political science can be carried out either inductively or deductively. According to Gerardo Munck and Richard Snyder (2007), the overwhelming majority of research in comparative politics is inductive. The inductive approach to theory is one in which theory flows from the analysis of observed facts. In other words, theoretical generalizations are built on the basis of specific facts, usually the data analyzed by the researcher. Although both qualitative and quantitative researchers engage in inductive analysis, game theoretic formal modelers of rational choice theory typically do not. Whichever method is used, inductive research typically contributes to generating new theories by specifying concepts and variables or by introducing new hypotheses to be tested. Inductive research is also particularly useful for studying areas of knowledge about which little is known and topics that lack a well-developed conceptual vocabulary. Comparative relationships between religion and the state are one such area of research. Jonathan Fox and Todd Sandler (2003) approach this issue area from the quantitative tradition in their article "Quantifying Religion," which develops a series of variables for measuring religion in comparative studies. In this case, the notion of *variable* is roughly equivalent to the *concept* that would result from similarly inductive qualitative work. Such concepts and variables provide essential components for deductive theorizing.

Deductive research begins with a theoretically derived hypothesis (King, Keohane, & Verba, 1994). As with the inductive approach, deductive theorizing is used by quantitative and qualitative researchers alike; it also forms the sturdy basis on which rational choice game theorists model action. A deductive approach to theory builds on a discipline's collective knowledge about a subject by encouraging researchers to form specific, testable hypotheses deduced from theoretical maxims and to submit those hypotheses to empirical tests. As such, the principal benefit of deductive research is its claim to produce cumulative knowledge. Another important benefit is the simple and powerful process that deductive theory generation prescribes for the conduct of research. Deductive reasoning requires researchers to deduce specific, observable implications of broad-gauged theories. In that way, it allows comparativists to address the most enduring questions in the field by using relatively little data (Geddes, 2003, offers a step-by-step procedure for formulating such questions). A potential weakness of the deductive approach is that it assumes that researchers have already amassed a great deal of coherent theoretical knowledge on a given topic. Indeed, whereas inductive reasoning, in its search for ever greater detail, risks infinitely delaying theory development, so deductive reasoning assumes that much of the theorizing has already been done.

Goals of Research

Although comparativists are united around their aim to explain and understand political phenomena around the world, their choice of method constrains them in the types of arguments they can make. *Designing Social Inquiry*, by Gary King, Robert Keohane, and Sidney Verba (1994), the most influential statement of the quantitative approach in the field, sums up the goal of research in a single word: *inference*. Inference allows researchers to extend their findings to other situations not directly observed by the initial study. In order to improve theory, King et al. outline a systematic, scientific procedure for testing theory aimed at producing valid descriptive and, preferably, causal inferences. A related goal of the quantitative approach is to maximize the researchers' leverage in explaining the phenomena of interest by allowing researchers to use the least amount of data to make the broadest generalization possible. While the authors of *Designing Social Inquiry* contend that their approach is suitable for both quantitative and qualitative work, most scholars within the qualitative tradition take a different view.

Because qualitative research has the largest, most variegated literature, as well as a plethora of distinct methodological tools, its theoretical goals are somewhat more diffuse. However, it is often said that whereas quantitative researchers are primarily concerned with explaining, qualitative researchers seek to understand. Although many qualitative methods seek causal explanations, practitioners in this tradition are more likely to be concerned with

understanding *how* a phenomenon came about than with explaining why it did. In other words, they tend to be more concerned with process than with probability or prediction. Charles Ragin, who has developed some of the most enduring qualitative tools (see, e.g., Ragin, 1987, 2000), describes the interpretive goals of qualitative research as "making sense of cases, selected because they are substantively or theoretically important" (Ragin, 2004, p. 109). Indeed, the pursuit of historical nuance and detailed narrative explain the tendency of qualitative researchers to focus on a small number of cases.

Whereas quantitative researchers seek to explain and qualitative researchers to understand, game theoretic modelers of rational choice theory aim their analysis at simplifying complex processes in order to predict. Rational choice-driven game theory is an individual-level theory that assumes that individuals attempt to maximize their utility, that decisions are made at points of equilibrium when "players" cannot increase their utility by making an additional move, and that the rules of the game are exogenous to the game itself (see Munck, 2001). Because these three conditions are assumed to be universal aspects of individual behavior, game theory purports to be applicable to *any* substantive question and able, therefore, to produce cumulative knowledge (for an important critique of the use of game theory in political science, see Green & Shapiro, 1994). While game theory is not the only framework used to carry out formal work in political science, it is by far the most common. Another formal approach is *network analysis*, which, although not as common in comparative politics, has already contributed to some substantive areas in the field and is poised to become an increasingly important method in the coming years (see Gould, 2003).

Methods of Analysis

Comparative methods mediate the scholarly encounter with observable facts by providing researchers with tools for analyzing data. Quantitative, qualitative, and formal methodological tools are differentiated by how they limit the scope of their research, how they measure the relevant variables or case aspects, and how they assess the theories they engage with.

Scope

Scope refers to the plausible applicability of a theory to a defined group of political situations or cases. In other words, the scope of a project informs its readers as to *what* precisely the research claims to create knowledge about and the relevance of its findings to other contexts and cases. Although many comparativists are concerned with the same "big questions," they disagree about which types of evidence should be employed to theorize about such questions. Thus, scope is the aspect of theory most closely related to data collection and analysis and depends in large part on the choices that a researcher makes in this regard.

The scope of a quantitative research project involves specifying the statistical model to be used, including the independent and dependent variables, and the number and quality of cases to be studied. It should be mentioned here that statistical models, which some consider formal (see King, 1989), are distinguished from game theoretic formal models of rational choice theory by the fact that variables in statistical models are typically closer representations of observable phenomena (see also Morton, 1991, p. 61). In terms of case selection, research norms in the quantitative tradition encourage the consideration of the entire universe of cases relevant to the phenomenon under study. What should be considered a case depends on the hypothesis and the unit for which it predicts outcomes. Thus, *case* may refer to a variety of units of analysis (i.e., state, party, city) or an event (i.e., civil war, policy selection, regime change). When it is not possible for a researcher to study the entire universe of cases, a sample from the universe should be taken in accordance with some substantive aspect of the theory (i.e., a given period), preferably at random, and in no case by selecting on the value of the dependent variable.

Certainly, choosing cases that have all experienced a similar dependent outcome in order to explain that very outcome leads to theoretical distortion in quantitative tests of theory. Yet resisting the temptation is not always intuitive. In fact, if one wants to explain why some states experience rapid economic growth in the wake of revolutions, it might seem logical to focus first on those cases in which such growth is known to have occurred, and only then attempt to explain what differentiates these cases from others. This would be a logical sequence for a qualitative researcher interested in developing in-depth knowledge of anomalous political processes or counterintuitive cases. However, if the researcher is more concerned with testing for the impact of theoretically relevant factors on the dependent variable, a strategy that begins with the universe of all cases would be a better fit. In fact, what distinguishes these research strategies from each other is not the absolute quality of the research involved but rather the scope of the arguments made possible by different types of research design (see Geddes, 2003, Chapter 3, for a more thorough analysis of this problem and its implications for comparative work).

Another important problem confronting quantitative researchers is the problem of *indeterminacy*. Indeterminacy usually springs from two sources related to specification of the model. The first is referred to as the *many variables, small N problem* identified by Lijphart (1971) and others. This problem arises when the number of inferences implied by a statistical model exceeds the number of cases. In such research designs, the number of cases could not possibly test for the causes suggested by the theory. The second most common reason for indeterminacy is *multicollinearity*. This problem arises when the explanatory variables of a statistical model are not independent of each other. For example, a study that seeks to explain the level of political

participation by women in new democracies might include variables measuring women's levels of education and women's workforce participation. To the extent that variation in the value of one of these variables predicts variation in the other, it would not be possible to measure the independent impact of either of them on the level of women's political participation in a given country. Statistically, problems of multicollinearity can be partially offset by increasing the number of observations. Such a strategy, however, runs the risks of either comparing cases that are not analytically equivalent or, if undertaken in an ad hoc fashion, altering the model without reference to theory. Despite these limitations, quantitative comparison has proven to be a useful and efficient method for testing hypotheses on large amounts of data that would be difficult to consider otherwise.

Scope is the most readily apparent difference between quantitative and qualitative work in comparative politics. While statistical work requires a relatively large number of cases, or observations, qualitative work tends to focus on a small number of cases. Part of this difference is semantic and attributable to the fact that the research questions of comparativists are often formulated at the level of the state. Even when the state is not the relevant focus of research, there is a substantial difference between the quantitative conception of a case as an analytically homogeneous unit among others and the qualitative view of a case as a "class of events" (George & Bennett, 2005, p. 17).

The scope of a qualitative research design ultimately depends on the goals of the researcher. If researchers aim to revise an existing theory or extend it, they will likely look to the literature for an anomalous case that has some potential to engage with the theoretical lacunae they seek to address. On the other hand, if researchers are interested in assessing the credibility of a theory, they might select a number of cases known to have experienced a similar outcome but whose histories they suspect involved different causal processes. This manner of case selection is starkly different from a statistical approach that warns against the analytical pitfalls of choosing cases on the value of the dependent variable. In cases of political phenomena about which there is relatively little theoretical knowledge, a qualitative research design may not be able to specify initially the cases under study. Such research designs, usually aimed at conceptual development or the construction of explanatory typologies, typically consist of a constant dialogue between theory and data aimed at understanding how to delimit the case itself and explaining what it is a case of.

The scope of a formal model rests on its assumptions and on how the model is constructed. As stated above, game theoretic models of rational choice theory assume that individuals seek to maximize their utility, that decisions are made at equilibria based on actors' preferences, and that the rules of the game are exogenous to the game itself. Because these assumptions are generally seen as universal, formal modelers of rational choice theory must use

some other criteria to explain their choice of scope. Indeed, rational choice theory does not itself stipulate any specific procedure for constructing formal models, and researchers in this tradition have not emphasized case selection as an important point of methodological reflection. Thus, during the late 1980s and 1990s, when game theory began to be used with greater frequency in studying comparative politics, the universality of rational choice assumptions became a subject of intense debate. In response, some researchers sought to limit the scope of rational choice theory either by relaxing its assumptions or by limiting its application to those cases in which its assumptions are most likely to reflect actual behavior. George Tsebelis (1990), for example, set forth the idea that rationality was a subset of human behavior more likely to describe situations in which the “actors’ identity and goals are established and the rules of the interaction are precise and known to the interacting agent” (p. 32). Yet others argued that much as regression analysis has, by necessity, an error term that provides researchers greater control in estimating causality, so formal models of rational choice theory are built on some false assumptions that facilitate hypothesis generation. Indeed, it is the simplicity of rational choice assumptions that allows the models to make clear and precise prediction. The more these assumptions are relaxed, the more difficult the model becomes to solve, and the less clear its predictions. In sum, the arguments that result from formal studies are relevant only to cases that fit the assumptions on which the model is based. Empirical work, on the other hand, is far more reliant on the precision of its definitions in specifying those cases to which its arguments can and cannot apply.

Measurement

Another area in which methods mediate the encounter between the researcher and the data is in measuring the concepts and variables used in a study. In every methodological tradition, researchers use measurements based on the goals of the research, the theory it engages with, and the requirements of their method. Researchers working in different methodological traditions typically have distinct vocabularies to describe their endeavors, and they often use different indicators to measure a concept labeled with the same word but having different meanings. Despite these differences, all comparativists strive for, and often claim to have achieved, measurement validity (see Adcock & Collier, 2001).

Comparativists often describe measurement in terms of levels. Scholars in the quantitative tradition sometimes distinguish their tradition from the qualitative tradition by their use of ordinal- and interval-level data and argue for the superiority of such measures while discounting the value of nominal data such as those used to create typologies. The claim of superiority of higher levels of measurement is based on the ability of statistical researchers to draw fine-gauged distinctions between large numbers of cases. However, qualitative researchers would argue that

such benefits are offset by the uncertainty of fit between such measurements and observed facts. Furthermore, Mahoney (2003), writing in the qualitative tradition, argues that the use of nominal and ordinal measurement is also central to the comparative historical approach and can be put to good use in determining necessary and sufficient causality in small-*N* studies.

While some of this disagreement is in fact substantive, part of it has to do with the relationship between measurement and the goals of research. For researchers in the quantitative traditions who seek to explain the impact of variables on an outcome, statistical models require measures that emphasize control. Furthermore, because such models usually test hypotheses on a large number of cases, researchers must use measures that can realistically be obtained in a fairly consistent manner for each case. Qualitative research designs, on the other hand, emphasize the credibility of measures for each case. Researchers in this tradition are more likely to develop highly nuanced measures of complicated variables, which accurately fit observations about the small number of cases considered. Indeed, in some qualitative research designs, the measurement of concepts may be the goal of the entire research project. Rather than measuring specific variables, formal modelers who use game theory must specify the components of their model, which usually include the relevant actors, their preferences and strategies, the level of information available to the actors, and the possible outcomes of the game. Although game theory does not recommend any specific procedure for conceptualizing a model, it rests on a well-defined set of universal assumptions that guide researchers in deducing these specifications from theory. Nevertheless, the absence of a single method for such an important aspect of modeling means that game theorists must rely on criteria exogenous to the theory itself. Although this encourages multimethod approaches, it introduces an element of potential inconsistency in the overall research design.

Theory Assessment

Given the variety of methods for generating theory, disparate goals of research, and logically distinct methods of data analysis, it is no surprise that different comparative methods also entail different ways of assessing theory. Indeed, both quantitative and qualitative methods mediate the dialogue between theory and fact. But whereas quantitative researchers tend to see a research project as one controlled communication, qualitative researchers are more likely to see the dialogue as a constant back-and-forth between theory and fact. Meanwhile, formal modelers of rational choice theory seek to contribute to theory by modeling the logical implications of its assumptions. These differing views of the nature of research directly impact how scholars use different comparative methods to assess theory.

The quantitative approach usually relies on a single data set to test the observable implications of theory in order to falsify or confirm it. For this reason, quantitative

researchers tend to design studies that rely on a large number of aggregated cases to observe the impact of independent variables on certain outcomes. Such large-*N* studies tend to assume a constant linear notion of causality. That is, they assume that the effects of independent variables on dependent variables are constant for the episode under study and that the causal impact is direct. They further assume that the outcome in one case does not impact the outcome in other cases. In sum, quantitative researchers take a counterfactual view of causality. One way to imagine counterfactual causality is by positing two parallel universes in which everything is the same except the value of a researcher's independent variable that alone explains the presence or absence of a given outcome. Of course, in observational studies, these universes do not exist, so causal inference must make up the gap. By accepting a counterfactual view of causality, quantitative work strives to approximate experimental work. In the absence of the perfectly controlled parallel universe required to carry out experimental research, quantitative analysts use statistical controls to decrease bias and improve the quality of inferences made from observational data.

In the constant dialogue between theory and fact that qualitative researchers undertake, it would likely be impossible to use new data for each encounter with theory. Because qualitative researchers are not generally constrained in their research by the controls of experimental logic, they can use the same data to test and refine their hypotheses. Thus, qualitative research designs tend to favor theory assessment over testing.

One method qualitative analysts use to assess theory is what is known as the *congruence method*. According to Alexander George and Andrew Bennett (2005), the congruence method is one in which a researcher “begins with a theory and then attempts to assess its ability to explain or predict the outcome in a particular case” (p. 181). Thus, it assesses the degree to which there is a fit between a theory's hypothesized causes and a case's observable outcomes. Among the advantages of this approach is that it can assess the ability of more than one theory to explain a given outcome. This is particularly important because it addresses the problem of *equifinality*—that is, that a single outcome may have multiple and unrelated causal paths. But because the congruence method, like many statistical methods, cannot explain why some theories are more congruent with outcomes, this approach is most usefully combined with other qualitative approaches that are more process oriented.

Qualitative researchers have not limited themselves to theory assessment but also seek to test theories using a variety of methods. It is important to point out, however, that a qualitative approach to theory testing differs substantially from quantitative, control-based theory testing focused on falsification. Bennett (2004) describes the goal of what he calls the “mechanism model of theory testing” as “to expand or narrow the scope conditions of contending theories as the evidence demands, and to

identify the conditions under which the particular causal mechanisms hypothesized by these theories interact with one another in specified ways” (p. 50). Such an approach is particularly well suited for addressing the equifinality problem and answering the “how” questions that qualitative researchers tend to ask. It also helps researchers understand why multiple theories are feasible because it can demonstrate how mechanisms from different theories interact with one another.

It should be noted, however, that the causal claims of such a theory rest on a distinct notion of causality that has important implications for how theory is assessed. Quantitative researchers using statistics usually rely on *probabilistic causation*, which assumes that every observable occurrence in the world is the result of at least some random causes that the research is unable to specify. Qualitative researchers, on the other hand, tend to see causality as more deterministic, assuming that every occurrence in the world is fully explicable because it is the result of some prior occurrences. The latter view explains why many qualitative researchers focus on identifying necessary and sufficient causes by specifying the conditions under which a particular phenomenon occurs. These differing views of causality also explain why qualitative researchers may choose to examine anomalous cases, logically positing that if a general theory does not fit for a specific case, then it must be revised. Although most researchers in either tradition are not likely to fully endorse either view, such assumptions about causality are implicit in the methods that researchers choose, and they limit the conclusions that researchers can reach (see Mahoney, 2003).

As mentioned above, the solution of formal models does not in itself constitute an assessment of the theory being modeled; rather, it presents a formal simplification of it. The major output of formal research, then, is not a clear assessment of theory but a set of hypotheses to be tested using a different methodology. Munck (2001) states the situation as follows:

Though models are ultimately assessed in terms of the empirically tested knowledge they generate, the exercise of modeling proper culminates in the proposal of hypotheses. Thereafter, modelers should test these hypotheses. But a formal methodology does not have direct implications for the testability of hypotheses; nor does it offer any guidelines about how to conduct the testing. (p. 200)

Indeed, game theory has been criticized as tending toward “pure theory” because its practitioners have rarely carried out the empirical evaluation their models call for. In response to these criticisms, and in the absence of a method for theory assessment internal to the method, some game theorists have made explicit efforts to lay the foundation for multimethod work. In *Analytic Narratives*, Robert Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast (1998) set forth a method

that combines formal modeling with qualitative analysis, while in *Methods and Models*, Rebecca Morton (1991) demonstrates how empirical statistical analysis can be used to test hypotheses derived from game theory.

Applications

The previous section outlined the ways in which methods mediate the researcher's encounter with theory and fact. An effort was made to show how the choice of methods mediates the scholarly encounter with theory and fact in terms of theory generation, the goals of research, methods of data analysis, and theory assessment. This section discusses three exemplary works in the field to demonstrate how these principles have worked in practice.

Qualitative

In a standard-setting work, Ruth Berins Collier and David Collier (1991) studied the process of *labor incorporation* in a paired comparison of eight Latin American countries: Brazil and Chile, Mexico and Venezuela, Uruguay and Colombia, and Argentina and Peru. These pairs represent what Przeworski and Teune (1970) would call "most different" systems, chosen on the basis of similar patterns of labor incorporation. By contrasting a comparably large number of cases, Collier and Collier highlight the significant differences between Latin American contexts while at the same time making an important theoretical and methodological contribution to comparative politics.

The Colliers situate their study in the literature on bureaucratic-authoritarian models that explain the collapse of democracy as a result of conflicts between workers and owners that arise as countries move from early industrialization to a more advanced economy requiring more intense capital accumulation to produce more sophisticated products. The Colliers critique this economically driven model by placing more emphasis on political factors. The basic argument they advance is that the process of labor incorporation in these states represents a critical juncture in the state's history that shapes legacies both in the short-term "aftermath" and in the long-term institutional "heritage" of a political system. Ultimately, it is these processes that explain why some states experienced the breakdown of their democratic systems whereas others remained more stable.

Their analysis, firmly within the tradition of historical institutionalism (see Thelen, 1999), begins with the emergence of a working class in each state. In nearly 900 pages, they develop a complex historical argument that can only be grossly simplified here. Using both within-case and between-case methods of analysis, they analyze the process of labor incorporation with a particular focus on labor groups, oligarchs, and reformers and the configuration of coalitions among them as they struggle for power. The

relative strength of the oligarchy is seen as particularly important. Whereas a weaker oligarchy provides greater coalitional space for reformers and leads to the mobilization of labor, a stronger oligarchy limits the political space open to reformers, who respond by seeking to control labor. It is important to note that in none of their cases does the working class initially emerge as autonomous, able to effectuate political change on its own. Rather the institutional configuration resulting from elite choices seemed to provide more or less space for labor activism in the aftermath and heritage phases of labor incorporation.

The main methodological contribution of this work is the concept of *critical junctures*. In their analysis, critical junctures are seen much as their ordinary language use would imply, that is, as pivotal moments that transform society and that have long-term effects. Labor incorporation is hypothesized to constitute such a critical juncture, developing along two dimensions, resulting in four patterns of labor incorporation: *radical populism*, *labor populism*, *electoral mobilization by a traditional party*, and *depoliticization and control*. Collier and Collier use historical comparison to test this hypothesis and find that it can at least partially explain the breakdown of democracy in Argentina, Brazil, Chile, and Uruguay and in every case demonstrates that labor incorporation had an important impact on events in the post-World War II era by shaping the political arena of the states under study. Thus, Collier and Collier's historical analysis represented an important theoretical innovation that ran contrary to most analyses of Latin American regimes. The potency of their analysis led many researchers to adopt and reuse their conceptualization of critical junctures as a way to make sense of slow-moving causal processes without reverting to a variable-oriented approach.

Quantitative

The relationship between economic development and democracy is one of the most contentious political issues that comparativists have consistently addressed in the past century. Przeworski and his colleagues Michael Alvarez, Jose Cheibub, and Fernando Limongi (2000) made an innovative contribution to this literature with their book *Democracy and Development*. The central question they address is, how do political regimes impact material well-being? To address this question, they use an inductive approach that gathers data on every country for which data were available for the period 1950 to 1990 and build an argument based on their findings at each step of the research.

First, they choose a minimalist definition of democracy suitable to their research question. Then they derive a set of rules that they use to define the cases in their universe as dictatorships and democracies. Using this descriptive data, they then use *probit analysis* to investigate the relationship between economic development, regime type, and survival. Using *lagged time series data*, Przeworski et al. (2000) then

consider the relationship between political regimes and economic growth. Here they mobilize their data to engage with the long-standing debate over whether democracy hinders economic growth by shifting resources from investment to consumption. After finding that political regime type does not impact economic growth, they turn to the question of political stability. From their exploration, they discover that instability means quite different things under different regime types and has a much greater impact on dictatorships than on democracies. In their final chapter, they investigate the paradox that population growth in dictatorships offsets higher rates of per capita income growth in the same states. Here their counterfactual statistical model leads to the striking conclusion that differences in a range of demographic indicators cannot be explained by exogenous factors but in fact stem from differences in the regime types, particularly the political uncertainty experienced by people living under dictatorships. Thus, each chapter of this study moves from a set of observations to a new set of questions, building a sophisticated statistical analysis, clearly outlined and explained in appendixes at the end of each chapter.

The inductive approach used by Przeworski et al. (2000), however, should not be seen as theory neutral. On the contrary, it is deeply engaged with existing theory, using previous analyses to guide the search. But their primary innovation is methodological. They suggest that most work done on the relationship between democracy and development is inconclusive because it is based on a counterfactual notion of causality but is not tested as such. By deliberately acknowledging the need for a counterfactual approach to causality in their statistical analysis, Przeworski et al. are able to arrive at new conclusions using data largely similar to that of other researchers before them. Among their most important findings is that democracies tend to have higher levels of economic development, not because development causes democracy, but because democracies are more likely to survive if the society is affluent. They also found that although democracies were particularly sensitive to economic crises, they were absolutely ensured of survival if they had reached a threshold level of per capita income. These theoretical contributions flow largely from the logical, explicit research design employed by the research team. In many ways, their study is not typical of quantitative studies in comparative politics. To begin with, they take an inductive approach to address a question that had previously been addressed by many other scholars. Furthermore, they use a series of statistical tests to assess hypotheses derived from an ongoing dialogue with theory that builds on the data being analyzed in the study. Their innovative approach, lucid writing style, and transparency of method have all contributed to this work's endurance in the field.

Formal Modeling

Josep Colomer's *Strategic Transitions* (2000) opens with a powerful and revealing statement: "Transition from

a nondemocratic regime by agreement between different political actors is a rational game" (p. 1). It is clear throughout his analysis that the model he creates is not meant as a metaphor for what happened when the Soviet Union dissolved but as an accurate, descriptive explanation. He does not say that transitions are *like* games but that they *are* games. The question his work addresses is, how is it possible for rationally motivated, self-interested actors to agree on transition? This is an important question, not only because it was historically surprising and unpredicted, but also because it is rare for such dramatic transformations to have taken place in such a short time with relatively little violence. After presenting a historical sketch of the historical background and the circumstances leading up to the fall of the Soviet Union, Colomer deduces the relevant actors and their strategies and preferences. The starting point of Colomer's analysis is that when an authoritarian regime is challenged, there are two possible outcomes: civil war or an agreed-on transition to democracy. In order to model this transition, Colomer uses the prisoner's-dilemma game as well as "mugging" games to identify equilibria. Most methodologists contend that game theory is best applied in situations in highly institutionalized settings such as parliaments or individual voting behavior. One of the innovations of Colomer's approach is that he applies game theory to a situation in which rules and institutional constraints are in flux. He justifies this approach by arguing that the outcomes are well defined and that in such situations, individuals are likely to make an important difference in the outcomes selected. Colomer contends that because the outcomes are known to the actors and because the actors are able to calculate that their choices would lead to suboptimal outcomes, they agree to some binding rules before engaging in the game.

In assessing the implications of this argument, Colomer's analysis draws heavily on the empirical record, but it does so primarily to buttress the argument rather than to test it systematically. He finds that transition by agreement is possible when (a) maximalist actors are weak, (b) the relevant actors are sufficiently strategically distant from one another, and (c) actors are farsighted enough to avoid strategies that result in myopic equilibria. The formal models analyzed are used to identify three models of transition, which he labels *transaction*, *negotiation*, and *collapse*. He then uses these models to explain the separation of the Soviet Union and the *Polish Roundtable*. In a final chapter, Colomer extends his analysis to show how the different models of transition impact institutional choice in the new post-Soviet states. Colomer's innovative methodology clearly achieves the objective of simplifying a complex set of strategic interactions. The delineation of the actors' preferences and strategies is valuable in itself, and the analytical exercise he presents, even if one is not convinced by the strong claim of descriptive explanation that he promises, adds enormously to the literature on democratization and remains an exemplary work of formal methodology.

Future Directions

This chapter has contrasted quantitative, qualitative, and formal approaches and has shown how they mediate the researcher's encounter with theory and fact. Empirical and formal comparative methods were presented, as well as the implications for research design of the three main methodological traditions.

Any such summary will nevertheless pass over the many ways in which researchers working in each of the traditions push and pull the field in different directions. To the extent that the choice of method flows from personal intuition or a well-reasoned belief about what exists in the world and how we learn about it, a researcher may be less flexible or less accepting of approaches that contradict a certain set of principles. Others may be driven by a particular political problem, making them more open to a variety of methodologies but less likely to give value to the generalizations that comparativists often make. Many researchers may also find themselves constrained by their own methodological training, unwilling or unable to invest in learning other methods, and as a consequence, they advocate certain traditions over others even when the traditions' shortcomings are clear. Thus, within each tradition, some researchers push for more methodological pluralism and others work within traditions, seeking hegemony over the research agenda of comparative politics as a field. Both positions can be fruitful and innovative, creating useful methodology syntheses or greater technical specificity for their approaches, but ultimately they have little to do with the ability of comparative research to explain or understand political problems.

Indeed, the community of comparative researchers is methodologically diverse, but its reasons for being so may have as much to do with theory and method as with larger social changes such as research funding, the structure of universities, the overall economic situation, and the quality and character of graduate education. Therefore, changes in these factors will have a great impact on the future direction of comparative politics. For example, publicly funded research projects may be more problem focused and require multidisciplinary team research. If universities rely more on such research grants, they may be more apt at some point to dismantle the traditional divisions of departments based on disciplines such as political science and sociology and replace them with a more research-center-based model in order to more effectively compete for funding. Such a move would dramatically change the character of graduate education and the methods that comparativists rely on to address political problems.

More substantively, one of the reasons that debates about methodology can be so intransigent is that the methods that a scholar chooses reflect assumptions about both ontology (what exists in the world) and epistemology (how people learn about what exists; see Hall, 2003). Quantitative, qualitative, and formal approaches all assume a positivist epistemology, which assumes that researchers

are capable of discovering political realities that exist independently of whether or how they are studied. Yet the positivist underpinnings of these methods have been highly criticized, particularly by constructivists and other critical theorists widely influential in other disciplines. Such approaches, often grouped together under the label of post-modernism or postpositivism, tend to be more reflexive about the role of the researcher and tend to blur the lines between research, theory, and practice. Nevertheless, while the positivist consensus in comparative research does not seem vulnerable to total collapse anytime soon, the post-positivist challenge may be one exciting venue for methodological innovation.

Conclusion

The lack of consensus regarding how to address the substantive questions relevant to the field leads some to question whether the field is maximizing its potential to contribute to the cumulative knowledge about politics across the globe in a systematic way. Some believe that greater consensus regarding methodological choices would lead to faster accumulation of knowledge and improved quality of research, whereas others believe that productive tensions among competing approaches lead to a best possible, if not ideal, outcome. This disagreement springs from questions regarding the purpose of the field and the goals of research.

It has not been possible in this short chapter to discuss the entire range of techniques, models, and games that quantitative, qualitative, and formal modelers use to carry out comparative work. Some of these techniques are dealt with in other chapters of this volume, and many more are described in the additional readings listed below. Nevertheless, an effort has been made here to describe what is at stake when researchers choose their methodology and to provide references to some of the more important methodological works in the field.

References and Further Readings

- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529–546.
- Alker, H. R., Jr. (1975). Polimetrics: Its descriptive foundations. In F. Greenstein & N. Polsby (Eds.), *Handbook of political science (Political science: Scope and theory)*. Menlo Park, CA: Addison Wesley.
- Bates, R., Greif, A., Levi, M., Rosenthal, J., & Weingast, B. (1998). *Analytic narratives*. Princeton, NJ: Princeton University Press.
- Bennett, G. (2004). Testing theories and explaining cases. In C. Ragin, J. Nagel, & P. White (Eds.), *Workshop on scientific foundations of qualitative research* (pp. 49–52). Washington, DC: National Science Foundation.

- Brady, H. E., & Collier, D. (Eds.). (2004). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham, MD: Rowman & Littlefield.
- Collier, D. (1993). The comparative method. In A. W. Finifter (Ed.), *Political science: The state of the discipline II* (pp. 105–120). Washington, DC: American Political Science Association.
- Collier, R. B., & Collier, D. (1991). *Shaping the political arena: Critical junctures, the labor movement, and regime dynamics in Latin America*. Princeton, NJ: Princeton University Press.
- Colomer, J. (2000). *Strategic transitions: Game theory and democratization*. Baltimore: Johns Hopkins University Press.
- Eckstein, H. (1975). Case study and theory in political science. In F. Greenstein & N. Polsby (Eds.), *Handbook of political science* (Vol. 7, pp. 79–138). Reading, MA: Addison Wesley.
- Fox, J., & Sandler, S. (2003). Quantifying religion: Toward building more effective ways of measuring religious influence on state level behavior. *Journal of Church and State*, 45(3), 559–588.
- Gates, S., & Humes, B. (1997). *Games, information, and politics: Applying game theoretic models to political science*. Ann Arbor: University of Michigan Press.
- Geddes, B. (2003). *Paradigms and sand castles: Theory building and research design in comparative politics*. Ann Arbor: University of Michigan Press.
- George, A., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge: MIT Press.
- Gould, R. (2003). Uses of network tools in comparative historical research. In J. Mahoney & D. Rueschemeyer (Eds.), *Comparative historical analysis in the social sciences* (pp. 241–269). New York: Cambridge University Press.
- Green, D., & Shapiro, I. (1994). *The pathologies of rational choice*. New Haven, CT: Yale University Press.
- Hall, P. (2003). Aligning ontology and methodology in comparative politics. In J. Mahoney & D. Rueschemeyer (Eds.), *Comparative historical analysis in the social sciences* (pp. 373–406). New York: Cambridge University Press.
- King, G. (1989). *Unifying political methodology*. New York: Cambridge University Press.
- King, G., Keohane, R., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Laitin, D. (2002). Comparative politics: The state of the subdiscipline. In I. Katznelson & H. Milner (Eds.), *Political science: The state of the discipline* (pp. 630–659). New York: W. W. Norton.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review*, 65(3), 682–693.
- Mahoney, J. (2003). Strategies of causal assessment in comparative historical analysis. In J. Mahoney & D. Rueschemeyer (Eds.), *Comparative historical analysis in the social sciences* (pp. 337–372). New York: Cambridge University Press.
- Mahoney, J., & Rueschemeyer, D. (Eds.). (2003). *Comparative historical analysis in the social sciences*. New York: Cambridge University Press.
- Morrow, J. (1994). *Game theory for political scientists*. Princeton, NJ: Princeton University Press.
- Morton, R. (1991). *Methods and models: A guide to the empirical analysis of formal models in political science*. New York: Cambridge University Press.
- Munck, G. (2001). Game theory and comparative politics. *World Politics*, 53, 173–204.
- Munck, G., & Snyder, R. (2007). Debating the direction of comparative politics: An analysis of leading journals. *Comparative Political Studies*, 40(1), 5–31.
- Przeworski, A., Alvarez, M., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development: Political institutions and well being in the world, 1950–1990*. New York: Cambridge University Press.
- Przeworski, A., & Teune, H. (1970). *The logic of comparative social inquiry*. New York: Wiley Interscience.
- Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. (2000). *Fuzzy set social science*. Chicago: Chicago University Press.
- Ragin, C. (2004). Combining qualitative and quantitative research. In C. Ragin, J. Nagel, & P. White (Eds.), *Workshop on scientific foundations of qualitative research* (pp. 49–52). Washington, DC: National Science Foundation.
- Sartori, G. (1970). Concept misformation in comparative politics. *American Political Science Review*, 64, 1033.
- Thelen, K. (1999). Historical institutionalism in comparative politics. *Annual Review of Political Science*, 2, 369–404.
- Tsebelis, G. (1990). *Nested games: Rational choice in comparative politics*. Berkeley: University of California Press.

35

CASE STUDIES

CATHERINE L. JOHNSON

University of North Texas

The case study method has always been an integral tool in the investigation of social science phenomena, being of particular value when the number of observations, or cases studied, is limited in number, restricting the utility of statistical approaches. However, for some time the individual case study approach had been supplanted by large-*N*, data-intensive quantitative methods as the preferred technique for empirical studies. More recently, the case study has seen a revival of interest by social scientists as part of a multimethod, holistic approach that includes formal, qualitative, and quantitative methods. Indeed, each major methodological approach plays an important role in the research cycle, with the qualitative application of the case study enlightening the inductive aspect of theory development through the identification of alternate causal explanations, new variables, or complex interactions of variables. Fundamentally, case studies allow one to go beyond often simplistic quantitative analysis and develop contextually rich and in-depth pictures of the phenomena being observed.

By itself, a case study is the history of an event, be it of short or long duration—a civil protest movement, for example, or the evolutionary process from colonial rule toward stable democracy. As such, a case study identifies the expected, predictable aspects of an event, while ideally it also captures additional but less quantifiable detail, such as the cultural context, that potentially asserts a causal role as well. Individual or comparative case studies of specific,

individual events, actors, or systems allow the researcher to obtain a depth of knowledge and understanding about the object being studied that large-*N* quantitative studies fail to provide.

A carefully crafted case study serves several purposes within the research cycle. First, while quantitative studies identify outlying or deviant cases, those well beyond the expected normal distribution, quantitative methods are generally not able to explain the specific reasons for a particular case's extreme variation from its population mean. The case study, however, not only provides the opportunity to identify likely reasons for these individual deviations but may illuminate previously unidentified causal variables and possible alternate explanations as well. This information potentially leads to the extension of existing theory, if not its revision, and may suggest new theoretical explanations altogether.

Additionally, the case study may be the best, or only, way to study certain phenomena because of the relatively small number of identified cases and a resulting scarcity of data, which restricts the use of quantitative methods. And while much of the earliest criticism of case studies (by social scientists) centered on their application as a mainly historical narrative, the substantive purpose of case study is to understand that history but to do so in a way that allows for the identification of critical actions, structures, or other aspects that contribute to the end result. Being able to examine with scientific rigor phenomena that either

do not lend themselves well to quantitative study, or for which only a limited set of objective measures is available, makes such an approach valuable. The role case studies can play in identifying and understanding previously unknown variables and in establishing causal paths and the interdependency of variables, as well as being critical tests of existing theory, makes them not just a complement to quantitative methods but potentially of equal value (Geddes, 1990; Gerring, 2004).

Case studies are by definition qualitative, meaning that the focus of the study is not primarily the systematic manipulation of aggregated points of data, an objective exercise, but rather a study that focuses on the quality of the potential data observed, a much more subjective work. This is not to say that case studies are not objective as well: In reality, for a case study to have any influence, it must identify and measure variables to allow for reliable comparison and to build theory that is testable, replicable, and generalizable. Case study is ultimately a method that falls into two forms: the *individual, within case study* and the *comparative across case study*, usually limited to a small number of cases. Both types work to identify causal relationships and enlighten theoretical explanations. Good case study work can be either *accumulating* (building on previous knowledge) or *original* (establishing entirely new avenues of research).

The Debate Within the Discipline

Political scientists have had an ongoing discussion about the role of the case study approach in their field. This discussion has focused on the relative value of case study compared with other methods for evaluating and advancing theoretical understanding. Of central concern is the perceived methodological limitation of single and small-*N* case work within a discipline that favors quantitative methodologies. A tension results between the benefits accrued from this method and its limitations. What value can a unique examination contribute? Are hypotheses and theory valid only if they are testable and generalizable? Within these debates over the fundamental usefulness of deliberative case study work are questions that address both the inherent strengths and weaknesses of such an approach. Scholars have generally fallen into two camps, those who argue for its usefulness and those who contend it has limited utility in a discipline with a strong quantitative emphasis and reliance on scientific method.

Addressing this fundamental question over the potentially ambiguous nature of a case study finding, which alone can neither directly inform nor disprove a generalizable finding, Arend Lijphart (1971) states that because of its singular nature, the case study in and of itself does not directly satisfy the standards of scientific research. He does, however, credit the case study with multiple *indirect* benefits, making it a valuable component in establishing political science theory. He identifies six types of case studies that fall into roughly two categories: those chosen because

the case itself is of interest and that are purely descriptive and those chosen to inform and build theory. The first category encompasses single case studies, which are generally detailed histories of a specific event or result and which, he argues, have value for this history alone. The thorough knowledge of a country gained by such an intensive, rich study provides critical information that others can also benefit from. Additionally, these in-depth analyses not only are a source of data for larger comparative studies but may also identify new variables of interest or suggest potentially new theoretical explanations. Lijphart's other typologies include those case studies that are chosen specifically for theory-building purposes. They include hypothesis-generating cases in areas in which no established theory exists; theory-confirming and -informing cases, both of which test existing theories; and deviant case analysis, for cases known to have varied from the expectations predicted by theory. This third type of case often reveals additional variables previously unidentified. It may suggest a temporal ordering of variables (path dependency) or identify the sometimes critical interactions of variables. The study of deviant cases may merely suggest refinements to the way variables are operationalized within the study, still an important theoretical contribution. These last three case typologies constitute the core of comparative case study, with their usefulness coming from their deliberate selection as a test to existing theory. While Lijphart identifies certain benefits of the case study approach, his praise is still conditional, and he favors the value of large-*N* quantitative studies whenever possible.

Harry Eckstein (1975) addresses the utility of case studies by first noting the predominant status held by historiographic work in earlier political science research. His main contention is that this early case study work, at both the micro and the macro level, although insightful in its own right, was perceived to be severely limited in its usefulness for producing generalizable theory, because of its singular focus and the statistical consequence of an *N* of 1. The prevailing assumption was that what theory-building utility case study work had was inductively drawn from the events studied, and those inferences might or might not represent replicable conclusions. Eckstein questions this assumption and lays out a detailed argument supporting the utility of case study work in all stages of the theory development process, not just the nascent ones. He additionally contends that case studies may actually be most valuable at the theory testing stage. Particularly in the field of comparative politics and when studying complex, potentially unique systems, Eckstein suggests that well-designed case study methods may be the best way of testing hypotheses and cumulating generalizable theories. Indeed, he emphasizes the role of case study in its comparative application and perspective.

Before he makes his argument for the value of case study to theory building, Eckstein (1975) provides valuable definitions of case study by emphasizing the concentrated, yet flexible, aspect of an investigation into a single event or individual. This focused yet not narrowly defined

approach allows the investigator to be open to unexpected observations and new conclusions. Eckstein additionally makes the important distinction that the study of one event does not necessarily mean only one measure of the results. Rather, he contends that how an event or thing is studied will dictate its number of observations. Thus, one event can be broken down into numerous observations. For example, “A study of six general elections in Britain may be, but need not be, an $N = 1$ study. It might also be an $N = 6$ study. It can also be an $N = 120,000,000$ study” (p. 85). This example illustrates his definition of a *case* as the single measurement of a pertinent variable observed, so that *comparative study* is then defined as “simply numerous cases along the same lines, with a view to reporting and interpreting numerous measures on the same variables of different ‘individuals’” (p. 85).

After he provides a useful review of the steps toward the development of theory, first the question or puzzle, followed by the formulation of a hypothesis and then a test, with the cycle likely repeating itself as refinements are made, Eckstein proceeds to describe five distinct varieties of the case study and identifies the particular uses each has. The first of these, the *configurative idiographic study*, is meant to be a comprehensive study of its target but one that allows for intuitive interpretation of the facts. By definition, *idiographic* is individualizing rather than nomographic or generalizing. Indeed, Eckstein acknowledges that this type of case study was the predominant type he first alluded to in this work. But he makes the point that the strengths of these types of case studies are their very weakness. Their rich description and often persuasive intuitive interpretations may be individually factual, but they aren’t systematic, which makes generalizable conclusions problematic and substantive theories unlikely.

The *disciplined configurative study*, a term Eckstein (1975) credits to Sidney Verba, turns this relationship around somewhat; rather than building theories on interpretations, interpretations should be driven by theory. This implies that the details of a case should either confirm or disprove a theory that ought to apply to it. The problem with this approach is, as Eckstein points out, its “discipline.” The strict and usually narrow application to a case of a hypothesized theory should either confirm or deny it. In essence, Eckstein suggests that this approach may be too restrictive. It may also lack the flexibility to accommodate more intricate relationships not already identified or suggested by existing theory. He also worries that interpretation of cases on an existing theory presumes that the theory itself is correct and suggests that existing theory, however valid, may “*compel* particular case interpretations” (p. 104, italics added) with its emphasis on generalizability at the expense of more individualized findings.

Eckstein’s (1975) third type is *heuristic case studies*, which are deliberate searches for discovery, often a result of trial and error. These are meant to be creative, stimulating the imagination of the researcher toward new ways of looking at a problem, focusing on broader, more generalizable

relationships. This discovery is incremental and is often developed in sequential studies as the new theory is further refined. The reason for heuristic case study is given rather succinctly by Eckstein: Theories do not arise from data alone but rather from the imagination of the researcher, after discerning puzzles and then patterns. Case studies, with their intensive analysis, increase the likelihood that these critical relationships will be found, particularly when they are carefully chosen to advance theory building. One caveat Eckstein offers on heuristic case studies is that they often produce too much—multiple explanations, too many variables, and a resulting complexity of interactions that are not only unwieldy but make generalization impossible.

Case studies are also used to probe the likelihood of proposed theories, a form that Eckstein (1975) calls *plausibility probes*. These are an intervening step before testing, to determine whether the expense of testing is warranted. Although the usefulness of such a study is limited to this end, and alone it cannot confirm a theory, it can, however, improve the prospects of testing, and for this reason it has value.

A more critical example of case study in theory building is the *crucial case study*. Eckstein (1975) confronts the dilemma of a single observation and the inability to correctly determine a statistical relationship on the basis of such limited information as a source of potential error for any theory based on it. The *inductive fallacy* is the error made when one derives a theory from only the observed (gathered) data, without further testing. The critical caveat is that one cannot test a theory with the same data used to originate the theory, and therefore another such example must be found. The *crucial case* is just such a test of a proposed theory. If all those variables deemed critical to a theory exist, then the results should be as predicted by the theory. Conversely, one can study a case similar in most respects, yet lacking in the hypothesized critical components, as a way of demonstrating that similar results did not result because the causal variable was missing. Although these *most likely* and *least likely case study* designs cannot absolutely confirm or deny theory, they are important tests of the likelihood of the theory and the correctness of the causal relationships being proposed.

Eckstein’s (1975) thorough typology and analysis of the case study method methodically crafts an argument for the benefits of case study work. These include the insight made possible by the rigorous, thorough inspection in a carefully crafted case study and its across-discipline utility in identifying new variables and new causal mechanisms leading to the generation of new theory. To accomplish this goal, Eckstein emphasizes that case study selection must be driven by theory, and not by interest or convenience.

Charles Ragin (1987), in *The Comparative Method*, devotes a chapter to the discussion of case-oriented comparative methods and addresses the likelihood that even the most meticulously performed case study is unlikely to

produce definitive explanations. However, identifying critical contextual facts may help determine the causal relationships underlying the observed phenomena. It is important to note that Ragin emphasizes the value that an intensive case study accrues to its researcher. Deep understanding of an event or case in its entirety, rather than merely knowing pieces of information, allows for more contextually rich comparison to other events. This richness can only enhance the reliability of the causal inferences drawn. Such depth of knowledge is likely limited to a small number of cases, and indeed this complexity is a constraint on the case study researcher. Case study is, as Ragin shows, a successful strategy for analyzing complex, multicausal events and at the same time still cohesively connecting them theoretically. He concludes with a nice summation of the strengths of the case study method: Case studies make possible the discovery of patterns of relationships and difference, with all deviations requiring an explanation, necessitating a thorough knowledge of the data. Since case study work does not rely on statistical probabilities such as frequency or distribution, a single case can be critical and can potentially prove or disprove a hypothesis. Case study work is holistic and requires a thorough understanding of the entire event, not just targeted aspects of it, and finally, case study encourages creative new ways of examining behavior and events. Particularly in the identification of complex interactions and the importance of context in understanding their role, Ragin makes the point that case studies provide a methodologically distinct approach.

In *Designing Social Inquiry*, Gary King, Robert O. Koehane, and Verba (1994) argue that the same level of testable, scientific rigor can be applied to qualitative work that quantitative scholars are able to use in their statistically based work; qualitative work includes, of course, case studies. King et al. focus on research design with an emphasis on the logic of inference, to use the facts that are known to learn about facts as yet unknown. This is then used to identify causal relationships and construct theories that can then be tested. King et al.'s emphasis on the latter stages of research design, producing theory that is testable and thus falsifiable, challenges case study researchers to think rigorously about their work, to recognize the similarities of quantitative and qualitative work with respect to empirical rigor, and to approach their work as such. King et al. argue that the primary way to do this is to see qualitative data more quantitatively, and to accomplish this from a practical standpoint, they advise maximizing the number of observations (from which measures are taken) whenever possible. At the same time, when adding an observation is not possible, they recommend summarizing on the outcome of interest instead, in order to avoid issues of *micronumerosity* (having more variables than observations). Echoing Eckstein (1975), King et al. remind us that the size of a case study, its *N*, is often determined by the level of analysis chosen: Is it one single event, several incidents within that event, or many more individual acts? In addition to

constructing a design that allows for multiple observations, the authors emphasize the requirement of designing theories that can be falsified (i.e., the null hypothesis can be tested). King et al. also address the importance of reducing the potential bias introduced through case selection. They emphasize the care with which cases must be chosen, as there must be "the possibility of at least some variation on the dependent variable" (p. 129). Other potential sources of selection bias they cite are investigator induced: choosing cases because data are available or because one has a particular interest in or understands the language, or the larger bias that often occurs when case selection is correlated with the dependent variable. In this instance, the process being studied has already been selected for over time, leaving as evidence only its most recent iteration and losing any obvious trace of what may have been critically important in the intervening stages.

Not all scholars implicitly agreed with the arguments made by King et al. (1994), and a lively review symposium in response to it appeared in the journal *American Political Science Review*. In it, Ronald Rogowski (1995) challenges King et al.'s concern with the testability of single-observation studies and relates three examples of just such single-case studies that do succeed under this limitation. He offers additional examples in response to their admonitions against dependent variable selection bias and comments that without deliberate selection based on a case's anomaly (its status as a statistical outlier), one of the core benefits of case study work would be lost. Rogowski sums up by emphasizing the importance of not losing the benefits of good qualitative work at the expense of increased quantifiability. In the same symposium, David Collier (1995) also takes issue with how King et al. address selection bias. However, although Collier generally concurs with their position, he argues for a bit more nuance when one is faced with some of the realities of the comparative method. Additionally he identifies the importance of valuing the context of research findings as more important perhaps than their generalizability, and he gently suggests that King et al. could be less rigid in their appraisal of qualitative methods.

Since case study is just that, an intensive examination of at least one item, how cases are selected is a fundamental issue. In comparative case studies, this issue is particularly relevant because small-*N* studies suggest that there exists more than one unique example of what is being examined and therefore a larger population to choose from. As a result, concerns over potential selection bias contribute prominently in discussions of the case study method. In "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics," Barbara Geddes (1990) addresses this issue by reexamining three prominent comparative studies. She neatly demonstrates how the potential error of case selection on the dependent variable can particularly impact results in small-*N* studies. Essentially a primer on selection bias, this article outlines the importance not only of identifying the most likely causal reasons some event occurred, but also of examining

the counterfactual as well. Geddes makes the point that by not providing a larger sample, selected randomly (rather than on the dependent variable) for testing the proposed relationship between cause and effect, one is really comparing only “the differences among the selected cases” (p. 132). She then shows how such an error can also occur in a path-dependent argument. In both examples, misleading findings resulted from researchers’ not expanding the population from which the targeted cases were drawn. Had they done so, they would have had a larger and likely more random sample to test. Geddes’s final example involves time-series studies and the determination of the appropriate end point of a case study. In this instance, she shows how changing the dates of a study would affect its results drastically, and she also makes the point that historical case studies are especially vulnerable to selection bias based on the time frames chosen for analysis.

The more recent discussion of case study work has focused increasingly on understanding the role of this method as part of a comprehensive research strategy. John Gerring (2004) emphasizes how, by failing to accommodate the bounded aspect of case work, most commonly used definitions for case study are inadequate. He offers the definition of case study as “an intensive study of a single unit for the purpose of understanding a larger class of (similar) units,” with units being “spatially bounded phenomena” (p. 342). This implies the study of a unique event or thing, at one point in time, with the goal of generalizability, which, he argues, provides a more theoretically useful interpretation. Gerring then provides a comprehensive discussion of the methodological ambiguities that occur in case studies and identifies six areas in which case studies are vulnerable. With these as a guide, he outlines the strengths and weaknesses of case (within-unit) study versus across-unit study. He notes that the case study method is more suited to descriptive inferences than to causal ones. It is a method that has a special affinity with intensive, focused studies rather than those that are extensive and broad. Case study is more likely to have high internal validity and weak external validity. It facilitates the defining of causal mechanisms, and not the testing of causal effects, performing better when causal mechanisms are deterministic instead of probabilistic. Finally, case studies are well suited to exploratory research but are limited in their uses for confirming hypotheses, yet they are preferred when across-case studies cannot provide adequate variance for the relationship being studied. With this enhanced clarity, and by situating case studies not apart from but as a complement to non-case methods, Gerring suggests that case study methods should be accepted as an equally worthy methodological approach by the entire discipline and that rather than favoring one method over another (often exclusively), scholars should use the method most suited to their question, their data, and their theory.

With Alexander George and Andrew Bennett’s (2005) *Case Studies and Theory Development in the Social*

Sciences, the debate within the discipline over case studies is brought up to date. Both authors have been longtime advocates of case study methods, and this latest work is a very thorough argument for the value of case study methods as part of a research strategy that includes both quantitative and qualitative methods as well as formal theory (Bennett, 2004, offers a chapter-length article distilled from this material, as well). George and Bennett disagree with King et al.’s (1994) contention that there can be only one “logic of Inference” (p. 11). George and Bennett discuss the relationship between case studies and the systematic building of theory. They compare the strengths and weaknesses of case studies and first identify four strengths, all areas in which statistical methods tend to be weak. These include concept validity, the potential for discovering new causal variables and deriving new hypotheses, a better understanding of the relationship between causal variables and possible path dependency, and the ability to identify or model the complex interactions of these variables. Weaknesses of case studies include the potential for introducing selection bias from the cases chosen and the inability to accurately measure the relative strength of an effect. Also, because of their single or very small number, case studies are relatively unique and not necessarily representative; cases chosen from a small pool may not necessarily be independent of one another, and they do not have a rich number of observations from which to judge the strength of associations between variables. George and Bennett advocate the use of the *structured, focused comparison*, which allows for the collection of data that can be systematically compared with other cases as well as accumulated. In this way, scientific rigor is added, and the utility of case methods is likely increased. The authors then outline the method of case study, from designing the research to executing the study and to drawing conclusions from the findings. In all steps, the role of theory is predominant: It drives the design and motivates the findings. In addition to being the definitive authority on case methods, George and Bennett present a compelling argument for using multiple methodological approaches in a research program. Not only do they show how qualitative and quantitative methods complement each other; they integrate formal modeling as well. This approach is gaining momentum in political science today, making a qualitative skill set not merely useful but necessary.

Examples of the Case Study Approach

U.S. Politics

In the field of U.S. politics, the classic example of a grounded, participant observer case study must be Alexis de Tocqueville’s (1835/2004) *Democracy in America*. Although most modern scholars of U.S. politics solve their *N* of 1 problem by focusing on the subunits of U.S.

government, using states or administrations, court terms or congressional voting records as their unit of analysis, Tocqueville analyzed the United States as a single entity. He drew his conclusion, that it is citizens' affinity for joining in and participating at all levels of civic life that strengthens democracy and enables it to flourish, from his personal observations as he extensively toured the country in the early 1800s. A more modern work in U.S. politics that is rooted in qualitative case study work is Richard Fenno's (1978) examination of congressional members, *Home Style: House Members in Their Districts*, in which he used extensive interviews and considerable time observing congressmen, both in Washington, D.C., and, critically, in their districts. This self-styled "soaking and poking" enabled a comprehensive, in-depth observation that allowed Fenno to identify the paradox of individual representatives' being very well-liked by their constituents at the same time as the institution of Congress is collectively viewed much more critically, and he explains much of the paradox by the personal relationships developed through district service. David Mayhew's (1974) *Congress: The Electoral Connection* also looks at the relationship between members of Congress and their constituents and is another example of a work based on inductive reasoning rooted in extensive in-depth participant observation. Mayhew finds that it is the incentive for reelection that motivates the individual behavior of both congressmen and the Congress. Through committee assignments, leadership positions, and vote trading (among other means), congressmen ensure their reelection chances. Mayhew suggests that with Congress motivated as a whole by mutual self-interest, it is no surprise that the structural arrangements of Congress, its organization of the leadership and committee system, have evolved to facilitate this behavior.

Comparative Politics

The case study method is used most extensively in the subfield of comparative politics. Using primarily small-N research designs, many significant works have been produced. Included among these is Barrington Moore's (1966) *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Moore examines five societies to compare their experiences with modernization and the economic revolution that ensues. He concludes that there are three likely outcomes, dependent on the country's social structure, and these in turn predict the likelihood of a successful transition to democracy or descent into dictatorship. The most well-known example of an intensive single-country case study must be Robert Putnam's (1993) *Making Democracy Work: Civic Traditions in Modern Italy*, also an excellent example of historiographic work. The subject of lively debate within the discipline, *social capital*, that is, the extent to which citizens are participatory and invested in their communities as a result of their civic relationships, was found by

Putnam to be a necessary component of a successful democratic society. Putnam argues that the associational experience of northern and central Italy developed interpersonal trust and fostered more democratic local governments, but the lack of similar groups in the south left them with less. Another such single-case work is Robert Bates's (1989) *Beyond the Miracle of the Market: The Political Economy of Agrarian Development in Kenya*. This work, which focuses on the intersection between land use, government institutions, and public policies, relies on a critical understanding of the economic, political, and cultural forces at work in Kenyan society. The complex interplay of economics and politics that Bates studies is only fully appreciated when the cultural context is included; the influence of tribal affiliations and Kenya's British colonial legacy are just two examples. These kinds of rich, multilayered observations and intimate knowledge of a society can be accomplished only with case study methods, with which Bates combines quantitative rigor as well.

International Relations

International relations scholars have also extensively used the case study method to selectively examine the actions of elite actors and organizations during critical events. Case study work is used to evaluate existing theory as well as propose alternate explanations to better understand the often complex motivations of and among nation-states. In *Essence of Decision: Explaining the Cuban Missile Crisis*, Graham T. Allison (1971) examines the Cuban missile crisis and, primarily through interviews, reconstructs the often conflicted decision-making process of all the major participants. To do this, he approaches the same event from the perspective of three different decisional-behavior models. These competing approaches are collectively used to illustrate the author's thesis: that despite internal pressures to the contrary, it was the actions and the decisions of the two leaders that successfully resolved the issue. Alexander George and Richard Smoke's (1974) *Deterrence in American Foreign Policy* is an example of a focused-comparison case study that examines 11 instances of the failure of U.S. deterrence policy. George and Smoke use process tracing to establish the causal explanation, which would not be possible without the depth of knowledge acquired in these case histories. In doing so, they critique existing theory and are able to offer a new, more dynamic, explanation. In another example of a focused-comparison study, Stephen M. Walt (1987), in *The Origins of Alliances*, looks at alliance formation and contrasts two distinct types: those made for mutual support to defend against a threat and those that are more opportunistic (or perhaps pragmatic), in which one aligns with the threat itself. Walt then explores the likely causes of these choices, looking specifically at shared ideology and the influences of foreign aid. His concentrated case study of states in the Middle East during a single period allows

him to develop the depth of knowledge necessary for such a study, in which data alone would be inadequate.

Conclusion

As the previous examples illustrate, case study work is applicable to a broad range of theoretical questions. Indeed, for many situations, a case study examination is the only way to rigorously examine an event. Case study can be used in either half of the research cycle: to deductively test the hypothesized research question or to inductively explore the results of empirical observations. It is also a valuable method for developing original theoretical insight, which can often form the basis of a research design using more statistically robust methods. Case study in and of itself serves a vital informative purpose as well, allowing in-depth appreciation of often nuanced yet critical conditions of the larger phenomenon being observed. Finally, the case study is increasingly being appreciated as a necessary component of comprehensive political science research today: Together with traditional quantitative methods that provide reliable statistical probabilities for a tightly focused view, and formal theory methods that produce more soft-focused or abstract explanations, case study work provides a necessary contribution by filling in the gaps, compensating for the inevitable shortcomings when formal and quantitative methods are applied to real-life questions and problems. Most critically, a well-crafted case study gives the researcher a level of knowledge and understanding of the matter being examined that no other method allows. This benefit alone justifies the application of case study methods to social science research today and in the future.

References and Further Readings

- Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.
- Bates, R. (1989). *Beyond the miracle of the market: The political economy of agrarian development in Kenya*. Cambridge, UK: Cambridge University Press.
- Bennett, A. (2004). Case study methods: Design, use and comparative advantages. In D. F. Sprinz & Y. Wolinsky Nahmias (Eds.), *Models, numbers and cases: Methods for studying international relations* (pp. 19–55). Ann Arbor: University of Michigan Press.
- Collier, D. (1995). Review: Translating quantitative methods for qualitative researchers: The case of selection bias. *American Political Science Review*, 89, 461–466.
- Collier, D., & Mahoney, J. (1996). Insights and pitfalls: Selection bias in qualitative research. *World Politics*, 49, 56–91.
- Eckstein, H. (1975). Case study and theory in political science. In F. Greenstein & N. Polsby (Eds.), *Handbook of political science: Vol. 7. Strategies of inquiry* (pp. 79–137). Reading, MA: Addison Wesley.
- Fenno, R. (1978). *Home style: House members in their districts*. Boston: Little, Brown.
- Geddes, B. (1990). How the cases you choose affect the answers you get: Selection bias in comparative politics. *Political Analysis*, 2, 131–150.
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge: MIT Press.
- George, A. L., & Smoke, R. (1974). *Deterrence in American foreign policy: Theory and practice*. New York: Columbia University Press.
- Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review*, 98, 341–354.
- King, G., Koehane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review*, 65(3), 682–693.
- Mayhew, D. R. (1974). *Congress: The electoral connection*. New Haven, CT: Yale University Press.
- Moore, B., Jr. (1966). *Social origins of dictatorship and democracy: Lord and peasant in the making of the modern world*. Boston: Beacon Press.
- Putnam, R. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Rogowski, R. (1995). Review: The role of theory and anomaly in social scientific inference. *American Political Science Review*, 89, 467–470.
- Tocqueville, A. de. (2004). *Democracy in America*. New York: Library of America. (Original work published 1835)
- Walt, S. M. (1987). *The origins of alliances*. Ithaca, NY: Cornell University Press.

PART III

INTERNATIONAL RELATIONS

36

HISTORY OF INTERNATIONAL RELATIONS

ANTHONY D. LOTT

St. Olaf College

As a field of study, international relations (IR) is a young discipline. Its genesis can be traced back to the period immediately following World War I. In the aftermath of the war, philanthropists, scholars, and diplomats in Europe and the United States sought an understanding of the causes of war and the means by which to promote international peace and security. At its core, the initial study of IR was both normative and empirical. Normative IR theory seeks to provide a set of values that policymakers, diplomats, and other actors should follow in order to better the human condition. Empirical IR theory seeks to explain the underlying causes of political events. Originally, IR had the normative desire to achieve pacific relations between states and an empirical concern with investigating the underlying causes of war and conflict.

With this narrow focus on interstate conflict, the original scholars in the field drew their theoretical insight from philosophy, history, law, and economics. Early scholars began a practice in IR of drawing on the philosophical works of Thucydides, Niccolo Machiavelli, Hugo Grotius, Thomas Hobbes, Jean-Jacques Rousseau, Immanuel Kant, and others in search of a proper understanding of the causes of war and the potential for peace. The study of past historical events was used to develop general principles that might be employed to resolve current and future conflicts. The growing importance of international law as a tool for states

would be used by IR scholars to frame theoretical approaches promoting peace and security. In the decades following World War I, departments of IR emerged in Great Britain, Switzerland, and the United States to train diplomats and policymakers and further the theoretical study of the discipline.

Although the discipline began by focusing on the causes of war and the potential for peace, the complexities of world politics and the emergence of globalizing forces throughout the 20th century expanded the scope of IR to include the study of human rights, migration, environmental cooperation, economic development, ethnic conflict, nationalism, terrorism, and international crime. Today, IR scholars have developed sophisticated theories and models in order to study an ever-expanding set of issues and concerns. Constituting one of the main subfields in political science, IR continues to demonstrate how political power defines this growing set of issues and concerns.

In the discussion that follows, the historical emergence and intellectual scope of the discipline are explored by examining the development of IR theory throughout the 20th century as well as the broadening list of empirical issues analyzed by IR scholars. Following this review, the future direction of IR is discussed. At the end of this chapter, a list of further readings is provided that introduces the reader to the themes introduced and the concepts explored.

Historical and Theoretical Developments

Liberal Beginnings

In 1919, a wealthy Welsh industrialist by the name of David Davies provided funds to the University of Wales at Aberystwyth for the purpose of studying international relations. After witnessing the carnage of World War I, Mr. Davies was intrigued by the ideals represented in the League of Nations and dedicated funds to endow the Woodrow Wilson Chair of International Politics with a belief that humankind could overcome war. Sir Alfred Zimmern, a British historian, became the first scholar of international politics when he accepted the post of Wilson Chair. His work is characteristic of early scholarship in IR and focuses on issues of economic interdependence and cooperation through international treaty law. Believing that scholars could make a difference in the world around them, Zimmern and other liberals of his time sought practical institutional solutions for the problems of conflict in the world. This focus on institutional solutions would come to dominate early discussions in the discipline and exemplifies liberal IR theory. He had an interest in and affinity for the League of Nations as a mechanism to prevent conflict and promote prosperity and peace among states. Many of the liberal IR scholars of the time, including Alfred Zimmern and Norman Angell, were active in League affairs and accepted the political position of contemporary leaders like Woodrow Wilson, who argued that self-determination for peoples and state membership in organizations like the League could create the foundation for international cooperation and the transcendence of war as a policy of the state. The pinnacle of liberal IR thinking that understands law as the basis for peace is the Kellogg–Briand Pact, an international treaty formally titled the Pact of Paris that outlawed war as a policy tool for states in the conduct of their foreign affairs. This treaty was signed by more than 60 states and exists today as a reminder of institutional attempts to transcend conflict through international law. By outlawing war among the signatories, the treaty established a legal basis for trying state actors who violated the provisions of the treaty. Further, the treaty provided a solid foundation for a set of international norms limiting the use of violence in international conflict and constraining the actions of states.

For early liberal IR scholars, the international community had the potential to use international organizations, international treaty law, and state diplomacy to solve problems. When constructed correctly, agreements negotiated by diplomats, written into law, and managed by proper organizations could resolve long-term international conflicts. Political challenges posed by World War II and the cold war would require certain refinements to liberalism in the 1970s, but the core belief in the possibility for change and the potential to overcome conflict still remains among liberal theorists. (For a more detailed discussion of liberalism, see Chapter 38, “Idealism and Liberalism.”)

Realist Critics

Throughout the interwar period, as the period from 1919 to 1939 would come to be called, scholars concerned with a focus on institutional mechanisms to overcome interstate violence challenged liberal IR theorists by emphasizing how enduring laws of power and the inevitable consequences of an international environment defined by a lack of a global government (anarchy) undermined institutional attempts to achieve peace. Exemplified in the classic E. H. Carr (1940) book *The Twenty Years' Crisis: 1919–1939*, realist IR theory focused on state concerns with security and the ever-present quest for power. There had been great concern among realists that liberal scholars and diplomats had a naive interpretation of international affairs and an idealistic faith in legal and institutional solutions as a means to solving potential conflicts. Accordingly, liberals had underestimated the potential for states to dismiss their legal (treaty) commitments and withdraw membership from international organizations when their national interest ran contrary to that law or organization.

Realists argued that scholars needed to develop a more sophisticated understanding of the causes of war. The normative desire to prevent war, although noble, undermined a clear understanding of its causes. As IR scholars, realists insisted that scholars seek a better understanding of what caused international violence in the first place. Writers such as E. H. Carr (1940), Hans Morgenthau (1948), John Herz (1950), and others labeled liberal IR scholarship “utopian” because of the liberal reliance on institutional solutions. These realists offered a vision of international politics where the potential for war required scholars and diplomats to mitigate its effects rather than seek its transcendence. According to realists, there was a set of conditions that prevented humankind from transcending war as policy. Human nature, often defined as a quest for power, and the anarchical environment limited the effectiveness of institutional solutions to prevent war. Reviewing the Kellogg–Briand Pact that liberals extolled as a sign of moral and political development in international affairs, realists noted that by the beginning of World War II, many of the pact’s signatories were occupied by, or at war with, other signatories.

Policymakers, realists argued, should recognize and internalize the important lessons of the Thirty Years’ War (1618–1648). The peace treaty that ended this European war established sovereignty as a defining principle of each state and required each state to defend against external aggression rather than rely on other states for their defense. Sovereignty, or the principle that states have control and autonomy over their physical territory and the citizens or subjects in that territory, would come to dominate realist scholarship. (For a more detailed discussion of realism, see Chapter 37, “Realism and Neorealism.”)

Marxist Alternatives

During this early period of theory development, a third approach to understanding the causes of war and the

mechanisms for peace was emerging as a critique of both liberal and realist IR theory. Emerging from the works of Karl Marx, Friedrich Engels, and others in the last half of the 19th century, Marxian scholars introduced a radical retelling of international relations. These scholars explored how particular class interests captured the power of the state and harnessed its foreign policies in order to promote their interests. This approach challenged both liberal and realist conceptions of the state as a neutral agent with regard to the citizens or subjects within. When understood to be in the interest of the propertied (or bourgeois) class, the state was engaged in a policy of expansion and imperialism. War, as state policy, could be understood as the means by which states would expand access to commodities and markets abroad. Moreover, opportunities for increased profits during war made it a profitable enterprise for the capitalist classes. Since the burden of battle was borne by the lower classes, Marxist IR scholars emphasized how war was the result of a particular economic system.

This radical approach to IR challenges liberalism and realism in two ways. First, as a moral critique, Marxism explores how capitalism, as an economic theory, undermines the human capacity for empathy. As a basis for the economic ordering of society, capitalism results in the exploitation of certain human beings and the alienation of all human beings. Once alienated, human beings become objects to be used just as the state might use any other weapon of war. Georg Lukacs (1971), a German philosopher writing in the early part of the 20th century, explores these moral criticisms of capitalism in *History and Class Consciousness*. His examination of human alienation has been used by subsequent Marxist IR scholars to explain how modern warfare dehumanizes people. Marxism also critiques the empirical rationale for war. Because capitalism requires that markets grow, war becomes a necessity. Capitalists must employ the state in war making in order to increase profits. V. I. Lenin (1916/1964), in his analysis of the causes of World War I, explores this issue in *Imperialism, The Highest Stage of Capitalism*. (For a more detailed discussion of Marxist ideas in IR, see Chapter 39, “Dependency and World-Systems.”)

Economic Interdependence and Global Security Challenges

After World War II, the historical trajectory of IR was altered by two significant factors in world politics. First, the emergence of a new international economic order, increasing global trade and financial flows among states, prompted scholars to adjust the mainstream theories of liberalism and realism. In 1944, policymakers of the Allied states met at Bretton Woods, New Hampshire, in the United States to negotiate institutional structures to manage the postwar global economy. At the conclusion of this international conference, the states in attendance agreed to create the World Bank (known originally as the International Bank for Reconstruction and Development), the International Monetary Fund (IMF), and the General Agreement on Tariffs

and Trade (which became the World Trade Organization in 1995). These institutions, and the norms of free trade, financial transparency, monetary stability, and economic integration that uphold these institutions, offered IR scholars additional variables to study in order to understand the causes of war and the potential for peace.

In a historical context, the development of these economic institutions demonstrated the importance of the liberal economic idea that global peace would be enhanced if states cooperated through trade and monetary policies. Often called neoliberals, scholars have explored how states in international relations create long-term cooperative arrangements that endure throughout the decades. Scholars such as Robert Keohane (1984) continue to study the implications of an increasingly global economic order. Their focus is on the complex web of governance rules. International governance occurs in conditions of anarchy, where government does not exist. However, even without formal government, neoliberals demonstrate how governance rules proliferate among the states in international relations and order their behavior. It is often the case that these governance rules proliferate because international regimes have been created to enhance the cooperation among states. The term *international regimes* refers to sets of principles, norms, rules, and decision-making procedures around which actors converge on a given issue area (Krasner, 1983). These regimes exist without the need of a formal government structure. Regimes function to provide a level of ordered and predictable governance among states in international society. An example of how regime cooperation has limited the spread of nuclear weapons follows.

In addition to scholarship on international governance and the importance of regimes, neoliberal scholars have employed the shared values that democratic states have in maintaining liberal economic conditions to study a separate peace that appears to develop among democratic states. Scholars such as Michael Doyle (1986) have explored this democratic peace hypothesis, arguing that sovereign states with market economies, limited government, civil rights, and representative government do not go to war with each other. This represents a direct theoretical challenge to realism. If neoliberals are correct and the type of government matters in terms of the potential for interstate peace, then the proliferation of democratic states should reduce the likelihood of war in the future. A world made up of democratic states could allow for the transcendence of interstate war as a policy possibility. (For a further discussion of these studies, see Chapters 38 and 39, “Complex Interdependence and Globalization” and “International Political Economy and Trade.”)

A second challenge to traditional IR theory emerged after World War II with the advent of nuclear weapons and the global security threat posed by U.S. and Soviet hostilities during the cold war. Previous security threats involved state aggression and the proper international response to that aggression. The threat posed by great-power nuclear weapons required scholars to imagine global nuclear annihilation. A deterrence strategy known as MAD, or mutually assured

destruction (for a more detailed discussion of this topic, see Chapter 44, “Deterrence Theory”), emerged among strategic studies scholars and influenced the national security strategies of both the United States and the Soviet Union. In the 1960s, IR theorists debated the relative stability of an international system in which nuclear weapons existed as a global threat. Many realists (Art & Waltz, 1971; McNamara, 1968; Newhouse, 1973; Schelling & Halperin, 1961) outlined the merits of a MAD environment where states would learn that use of nuclear weapons would result in their own demise. This, they argued, would create a level of stability in international affairs and minimize the likelihood of systemwide wars. Scholars in other traditions (Bennett, 1962; Clancy, 1961; Dyson, 1979) contended that the potential for accidents or the irrational actions of one individual who did not learn the lessons of MAD could place billions of lives in peril.

Although the theoretical and moral debates remain ongoing in IR theory, the presence of nuclear weapons in world politics has led to broad agreement among diplomats and policymakers that access to nuclear technology should be regulated at the international level. The international community has developed an intricate set of principles, rules, norms, and decision-making procedures to limit access to nuclear technology and minimize its proliferation beyond a small group of declared nuclear states. These components constitute the nuclear nonproliferation regime. Actors in this regime include declared nuclear powers, the United Nations Security Council, and the International Atomic Energy Agency. The regime is centered on a multilateral treaty, the Nuclear Non-Proliferation Treaty. Each of these components includes a set of principles, norms, rules, and decision-making procedures that guide the behavior of states on the issue of nuclear nonproliferation.

Regimes exist in all areas of international affairs, including human rights, security, the environment, trade, finance, and cultural preservation. The study of international regimes has become a central research area in IR. Regime analysis has emerged as a useful approach to understanding conflict and cooperation. Throughout the 1980s and 1990s, scholars in IR (Keohane, 1984; Krasner, 1983; Young, 1989, 1994) produced numerous works that furthered our understanding of and appreciation for international regimes. This literature helps explain how governance without government is possible and why international politics is most often ordered and predictable. Sophisticated theoretical studies of regimes provide a more comprehensive picture of international affairs than the earlier theoretical work conducted during the interwar period. Because regimes include multiple actors (such as states, international governmental organizations, nongovernmental organizations, and transnational corporations), their study provides theorists with a more detailed model of international affairs. In addition, because regimes involve institutional rules (like international law) and socially appropriate practices (like international norms), their study provides an opportunity for a more comprehensive approach to the study of continuity and change in world

politics. (For a more detailed discussion of regimes, see Chapter 51, “International Organizations and Regimes.”)

Alternative Challenges to Mainstream IR Theory

Although the inclusion of factors such as globalizing economic structures and the presence of nuclear weapons offers IR scholars a new set of factors to include in their studies of international affairs, the treatment of these and other issues by mainstream scholars in the field has been viewed as inadequate by scholars critical of a focus on states instead of individuals and national security instead of human security. Alternative voices emerged throughout the 1980s that sought to critique both the mainstream IR scholarship of realists and liberals and the foreign policies that they studied. Although these alternative voices do not represent a single theory or approach to the study of world politics, they share a common concern that the discipline of IR and the practice of international politics have relied on concepts such as *state sovereignty* and the *state system* at the expense of other concepts. This state-centric emphasis marginalizes a set of concerns that need to be explored further if theorists wish to provide compelling and comprehensive answers to current and future problems.

The state, what constitutes it, what the implications are of particular foreign and security policies pursued by it, and where its national interests come from have been left underanalyzed and unquestioned. These questions represent a different type of question than those posed by realist and liberal scholars. Mainstream questions focus on the international environment and leave the internal assumptions of the theories themselves unexamined. Questions that critique the assumptions within theories are termed *critical questions* and require theorists to reexamine and reconstruct the theoretical foundations of international affairs. Often, this means that IR theory needs to be reformulated in order to remain coherent.

This alternative manner of theorizing has had a profound influence on the IR discipline. Feminist, constructivist, and environmental scholars represent important challenges to the traditional study of IR. Each of the approaches is examined below. Note that although each approach is different in its focus and the critical question that it poses, all of the approaches are similar in that they challenge liberal and realist IR theory.

Feminist International Relations

By asking an alternative set of questions, feminist scholars (Carpenter, 2006; Enloe, 1989; Tickner, 1992, 2001) have been able to provide insight into gender issues that remain hidden by standard approaches in the discipline. The general focus of the discipline on war and economic affairs marginalized gender inequality. Feminist scholars in the 1970s argued that traditional gender roles in society undermined inclusion of women in international affairs. Divisions

of labor in both advanced industrial and traditional societies mandated that women remain in the private sphere while men participate in the public sphere. Because war and diplomacy were public acts, women—and the issues of most concern to them—would be discounted. Similarly, because the home was part of the private sphere, feminist concerns of family, education, health care, and children would be marginalized, and issues of state GDP and increased trade would be emphasized. In both cases, feminist IR scholars articulated a new set of questions to challenge mainstream IR scholarship.

Consider the following example. Both realist and liberal IR scholars accept the state as a necessary actor in international affairs and argue that its presence enhances the security of individuals by protecting them (collectively) against potential harm that exists in the international (or external) environment. Realists argue this by employing a concept like the *national interest*, and liberals emphasize this by employing a concept like *collective security* through international law. In either case, both theoretical approaches accept that the state is a central variable in the maintenance of international peace and security. Feminist IR scholarship challenges this assumption and questions whether the state might reinforce social structures that oppress and exploit particular groups. Domestically, IR theories that promote the idea that states protect the national interest and maintain national defenses are participating in a public debate about where to spend limited tax revenues collected by the state. Given a limited amount of state funds that can be spent on all public goods, this has the effect of steering money away from social programs that might be used to educate children, provide welfare and child care assistance, and promote health care for vulnerable groups. If public funds cannot be provided to supply these goods, the burden of supplying these goods often falls on women. Internationally, IR theories that emphasize issues such as balance of power and alliance structures or foreign direct investment and increased global trade are reinforcing a set of social structures that exploit women. In an important early critique of IR, Cynthia Enloe (1989) argues that mainstream IR theory neglected to study the social implications of cold war bases around the world. In *Bananas, Beaches, and Bases*, Enloe directs the attention of the reader away from a standard view of international politics as a struggle for power and security and toward an analysis of the implications of foreign military bases in third-world countries. Recognizing that these bases reinforce stereotypical views of masculinity and perpetuate the exploitation of women who work in and around military bases, Enloe challenges traditional assumptions of international politics. (For a more detailed discussion, see Chapter 41, “Feminist International Relations.”)

Constructivist International Relations

A second alternative challenge to traditional IR scholarship has emerged among scholars interested in challenging the origin of state interests. Constructivists (Hopf, 2002;

Katzenstein, 1996; Lapid & Kratochwil, 1996; Wendt, 1992) focus on the formation of national identity as a prerequisite for understanding and explaining national interests. Issues of identity and the norms that shape and constrain it remain hidden by mainstream approaches to IR that assume a given and predetermined national interest exists among all states. As with the feminists, these scholars ask a set of critical questions that requires a reexamination of traditional theories. The aim of constructivist IR scholarship is to challenge the underlying motivations that both liberal and realist scholars assume states have when conducting their foreign policies. By challenging the essence of these mainstream theories, constructivists are engaged in more than correcting a perceived flaw in IR scholarship; they are also engaged in reimagining the conduct of international affairs and allowing alternative interpretations of historical events to emerge.

Consider the following example. During the cold war, American and Soviet identities were based on a consideration of the other as an enemy. Each state had a negative perception of the other based on the qualities one possessed as distinct from what the other possessed. The United States perceived itself in positive terms because it upheld democratic values and political and civil rights. It perceived the Soviet Union in negative terms because it claimed the Soviet Union did not possess these traits. The Soviet Union perceived itself in positive terms because it was concerned with economic and social equality. It perceived the United States in negative terms because it claimed the United States did not possess these concerns. As a result of these identity constructions, each state determined the other to be an enemy and subsequently viewed the other's actions as hostile and threatening. Constructivists argue that this scenario is what is missing from the work of mainstream IR scholars when they seek an understanding of the national interest. Only by identifying how national identities are created can the interests that form from those identities be understood. The events of the cold war come to be seen as a set of identity performances that reinforce a self–other dynamic in international politics rather than the logical outcome of two states pursuing predetermined national interests. As one prominent constructivist, Alexander Wendt (1992), has stated, anarchy is what states make of it; it is not an enduring cause of war in itself.

Constructivist IR scholarship has become an important voice in understanding terrorism, ethnic conflict, and religious violence. Constructivists have developed detailed case studies exploring how the formations of particular identities among one group exclude membership for other groups. These studies point out that these identities do not cause war but do give rise to a self–other dichotomy that can be exploited by political entrepreneurs seeking power.

Environmental International Relations

A third alternative approach to understanding international relations requires scholars to reexamine the ability

of the state and the state system to solve pressing ecological problems that are transnational in scope and require cooperation among multiple actors. With the rise of national environmental movements in the United States, western Europe, and New Zealand in the 1960s and 1970s, the international community held its first global environmental conference in Stockholm, Sweden, in 1972. The United Nations Conference on the Human Environment was first proposed by Sweden in 1967 and was later supported by the United States. Scientists and policymakers were becoming increasingly concerned that economic activity in one region of the world was affecting the quality of the environment in other regions of the world.

As IR scholars turned their attention to environmental issues, it soon became apparent that the mainstream theoretical emphasis on states, state sovereignty, and the national interest was not an adequate approach to resolving the pressing problems associated with the transnational dimension of the environmental problems. Realism and liberalism were constrained by a state-centric understanding of international politics. The world map that defines both theories is political. The world is divided into states with clearly defined borders. Ecosystems and environmental pollution, however, do not respect state borders. Environmental IR theorists (Haas, 1990; Luterbacher & Sprinz, 2001; Newell, 2006) questioned the disciplinary focus on a political world map and sought to reimagine the map as physical in nature. Political solutions to environmental problems require states, nongovernmental organizations, scientific groups, multinational corporations, and others to cooperate in ways that realists and liberals may not emphasize. Unlike peace agreements after major wars or security alliances during times of peace, solutions to environmental problems usually require the cooperation of more than just state actors. For example, state participation in a security alliance requires the cooperation of key government agencies within a state (the foreign and defense ministries, the chief executive, and a legislative body) but does not require much in terms of changes to the behaviors of the average citizen. Solving transnational environmental pollution, however, might require international governmental organizations, state agencies, corporations, and citizen groups to be involved in changing individual behaviors. Moreover, environmental problems are often linked to economic issues. Solving environmental problems can require states to forego economic development plans and limit short-term economic gains for the sake of improved long-term environmental sustainability.

These challenges to traditional IR scholarship require theorists to construct alternative understandings of international relations. Scholars in this area of IR have researched how environmental scarcity can be a cause of war. Thomas Homer-Dixon (2001) argues that under certain conditions, environmental degradation can contribute to international conflict. Scholars have also examined how the international community has responded to environmental concerns. By examining the institutional structures created since the first

international conference in 1972, scholars such as Oran Young (1989) and Peter Haas (1990) have contributed to the field by including epistemic communities (or groups of scientists with a vision of the problems and potential solutions) and regimes into the study on environmental IR.

In the aftermath of the 1972 conference, the international community has been active in institutionally managing the international environment. The United Nations created the United Nations Environmental Programme and held a subsequent international conference in Rio de Janeiro (the United Nations Conference on the Environment and Development) in 1992. Broad international treaties to manage the oceans (United Nations Convention on the Law of the Sea III), air pollution (Convention on Long-Range Transboundary Air Pollution), the movement of hazardous waste (Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal), and global climate change (United Nations Framework Convention on Climate Change) have been negotiated. (For a detailed discussion, see Chapter 53, "International Environmental Politics.")

Future Directions

The discipline of IR is adapting to new challenges at the dawn of the 21st century. States are confronted with emerging security threats including terrorism, cyber crime, and ethnic conflict. Challenges posed by uneven development, poverty, inequality, and malnutrition undermine possible state-centric responses. Humanitarian crises caused by political violence, corruption, and environmental disasters require substantial cooperation among international actors. A growing awareness of ecological interdependence demands that practitioners, scholars, and ordinary citizens reconceptualize international politics. Many of these new challenges are caused by a process of globalization that has been occurring for centuries. Globalization manifests itself in many ways but is most often referred to as a shortening of time and space that allows human beings to interact more directly than in times past. With rapid changes in communications technologies and information systems, groups once limited by time and space play an increasingly important role in international politics. These nonstate actors challenge IR scholars to incorporate additional variables into more complex theories of world politics.

New Security Threats

Although terrorism is not a new issue in international politics, the globalizing forces that allowed for increased economic trade and wealth also allow terrorists to strike at larger targets. State-sponsored terrorism has been a concern among IR scholars for decades. New forms of terrorism involve nonstate terrorist groups with political grievances against states. Terrorism is generally defined as a premeditated, politically motivated violent act meant to cause fear

among noncombatants. Nonstate terrorist groups challenge states in two ways. First, terrorist groups undermine the political fabric of domestic societies by invoking fear among the populous and undermining the legitimacy of the state to maintain peace and security. Second, terrorist activities challenge the foundation of international society by compromising sovereignty. IR scholars have adapted mainstream IR theories to incorporate terrorist activities. Current analysis seeks to understand the rationality of terrorist organizations and the security responses that states make in order to minimize terrorism.

Cyber crime is another emerging security threat that IR scholars have begun to investigate. A growing amount of national and international commerce and communication takes place electronically. Disruptions to the electronic infrastructure of global commerce threaten national economies and undermine the welfare of societies. In addition, states must protect electronic databases and the classified information they contain. New directions in security studies have been developed to understand and account for the challenges that states face with regard to cyber crime.

Increasingly, conflict between groups involves intrastate ethnic conflict rather than interstate conflict. This represents a theoretical challenge to a discipline founded to transcend or mitigate interstate conflict. As the preceding discussion demonstrates, mainstream IR theories have focused on understanding international wars and promoting effective mechanisms for peace. IR scholars recognize the need to develop a much more sophisticated understanding of conflict that can incorporate both intra- and interstate dimensions of conflict. For instance, recent works by Robert Jackson (1990) and Mohammed Ayoob (1995) explore the internal dimensions of conflict and provide a sophisticated understanding as to how the complex state-making process creates certain states beset by internal conflict and strife. Moreover, these studies demonstrate how these states undermine regional stability. Future research in this area will be necessary in order to develop increasingly useful theoretical models to predict potential areas of conflict and employ international resources prior to their onset.

Development Strategies and Humanitarian Crises

In September 2000, member states of the United Nations adopted a set of millennium development goals to reduce poverty and to increase education, access to health care, and gender equality by 2015. These development goals provide evidence of the continued shift away from the traditional issue areas of international politics. Increasingly, states recognize the need to cooperate on a number of issues that were once considered internal or domestic issues. With the challenges posed by the new security threats and a growing awareness and appreciation for cosmopolitan values, state actors recognize the need to share development strategies and improve the human condition for all. This concern over the welfare of all human

beings and a broad interest in humanitarian responsibility challenges earlier normative concerns in IR. Recent studies in IR involving issues of economic development, poverty, inequality, malnutrition, and humanitarian crises suggest a new normative shift in the norms and values examined by IR scholars. These new values are enshrined in concepts like a *responsibility to protect* those individuals and groups in states who are not being protected by their own states. This departure from traditional understandings of state sovereignty and the principle of nonintervention suggests a new debate about what constitute appropriate sovereignty is currently emerging among practitioners and theorists.

Ecological Challenges

In response to the first global environmental issues in the 1970s, states developed complex institutional mechanisms to manage these problems. The persistence and proliferation of these problems has increased the need to further study cooperative strategies for managing them. Declining biodiversity, a looming energy crisis, and challenges to adequate food supplies are three key areas of environmental concern. However, the most difficult environmental problem to solve appears to be global climate change. Insufficient compliance with the Kyoto Protocol and the development demands of industrializing states such as China, India, Brazil, and Russia require states to resolve long-standing collective action problems in order to construct effective treaties for solving climate change. Collective action problems involve scenarios where the most rational actions taken by individual actors are suboptimal for achieving group success. That is, the best option for the group is not necessarily the best action for each individual member of that group (Olson, 1965). Global climate change is often perceived to be a classic collective action problem. IR scholars interested in this subject are seeking more sophisticated theoretical approaches to resolving climate change by invoking complex and varied incentive strategies to achieve cooperation (Luterbacher & Sprinz, 2001; Newell, 2006).

Conclusion

Although a young discipline, IR has developed increasingly sophisticated approaches to explaining international conflict and the myriad issues that have emerged over the past 100 years. The complexities of world politics and rapid globalization require contemporary IR scholars to investigate more complex issues than those who originally developed the discipline. Although mainstream theoretical approaches to the study of international politics are still important in the field today, alternative theoretical emphasis on gender, norms, and environmental interdependence require scholars to consider a set of important theoretical questions left unexamined by mainstream approaches. Further, new security, humanitarian, and ecological

challenges appear to undermine state-centric approaches in the discipline and require scholars to push the boundaries of the discipline in new directions.

References and Further Readings

- Angell, R. N. (1910). *The great illusion: A study of the relation of military power in nations to their economic and social advantage*. New York: Putnam.
- Art, R. A., & Waltz, K. N. (1971). *The use of force: International politics and foreign policy*. Boston: Little, Brown.
- Ayooob, M. (1995). *The third world security predicament: State making, regional conflict, and the international system*. Boulder, CO: Lynne Rienner.
- Bennett, J. C. (Ed.). (1962). *Nuclear weapons and the conflict of conscience*. New York: Scribner.
- Bull, H. (1977). *The anarchical society: A study of order in world politics*. New York: Columbia University Press.
- Carpenter, R. C. (2006). *Innocent women and children: Gender, norms, and the protection of civilians*. Aldershot, UK: Ashgate.
- Carr, E. H. (1940). *The twenty years' crisis, 1919-1939: An introduction to the study of international relations*. London: Macmillan.
- Clancy, W. (Ed.). (1961). *The moral dilemma of nuclear weapons*. New York: Council on Religion and International Affairs.
- Doyle, M. W. (1986). Liberalism and world politics. *American Political Science Review*, 80(4), 1151-1169.
- Dyson, F. (1979). *Disturbing the universe*. New York: Harper & Row.
- Enloe, C. (1989). *Bananas, beaches, and bases: Making feminist sense of international politics*. London: Pandora Books.
- Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization*, 52, 887-918.
- Haas, P. M. (1990). *Saving the Mediterranean*. New York: Columbia University Press.
- Herz, J. H. (1950). Idealist internationalism and the security dilemma. *World Politics*, 2(2), 157-180.
- Homer Dixon, T. F. (2001). *Environment, scarcity, and violence*. Princeton, NJ: Princeton University Press.
- Hopf, T. (2002). *Social constructions of international politics: Identities and foreign policies, Moscow, 1955 and 1999*. Ithaca, NY: Cornell University Press.
- Hurrell, A. (2007). *On global order: Power, values, and the constitution of international society*. Oxford, UK: Oxford University Press.
- Jackson, R. (1990). *Quasi states: Sovereignty, international relations and the third world*. Cambridge, UK: Cambridge University Press.
- Katzenstein, P. J. (Ed.). (1996). *The culture of national security: Norms and identity in world politics*. New York: Columbia University Press.
- Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton, NJ: Princeton University Press.
- Krasner, S. D. (Ed.). (1983). *International regimes*. Ithaca, NY: Cornell University Press.
- Lapid, Y., & Kratochwil, F. (Eds.). (1996). *The return of culture and identity in IR theory*. Boulder, CO: Lynne Rienner.
- Lenin, V. I. (1964). *Imperialism, the highest stage of capitalism*. Beijing, China: Foreign Languages Press. (Original work published 1916)
- Lukacs, G. (1971). *History and class consciousness: Studies in Marxist dialectics*. Cambridge: MIT Press.
- Luterbacher, U., & Sprinz, D. F. (Eds.). (2001). *International relations and global climate change*. Cambridge: MIT Press.
- McNamara, R. S. (1968). *The essence of security*. New York: Harper & Row.
- Morgenthau, H. J. (1948). *Politics among nations: The struggle for power and peace*. New York: Knopf.
- Newell, P. (2006). *Climate for change: Non state actors and the global politics of the greenhouse*. Cambridge, UK: Cambridge University Press.
- Newhouse, J. (1973). *Cold dawn: The story of SALT*. New York: Holt, Rinehart & Winston.
- Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge, UK: Cambridge University Press.
- Sagan, S. D., & Waltz, K. N. (1995). *The spread of nuclear weapons: A debate*. New York: W. W. Norton.
- Schelling, T. C. (1960). *The strategy of conflict*. Oxford, UK: Oxford University Press.
- Schelling, T. C., & Halperin, M. H. (1961). *Strategy and arms control*. New York: Twentieth Century Fund.
- Tickner, J. A. (1992). *Gender in international relations: Feminist perspectives on achieving global security*. New York: Columbia University Press.
- Tickner, J. A. (2001). *Gendering world politics: Issues and approaches in the post cold war era*. New York: Columbia University Press.
- Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 46, 391-425.
- Wendt, A. (1999). *Social theory of international politics*. Cambridge, UK: Cambridge University Press.
- Young, O. R. (1989). *International cooperation: Building regimes for natural resources and the environment*. Ithaca, NY: Cornell University Press.
- Young, O. R. (1994). *International governance: Protecting the environment in a stateless society*. Ithaca, NY: Cornell University Press.
- Zimmern, A. E. (1939). *The League of Nations and the rule of law, 1918-1935*. London: Macmillan.

37

REALISM AND NEOREALISM

SVETLA BEN-ITZHAK

University of Kansas

Realism has long been one of the main theoretical approaches to the study of international relations. It is an intellectual tradition built on distinct concepts and arguments about what governs politics among states. As such, its main precepts assert that the international system is characterized by anarchy, states are its principal actors, which are sovereign and rational acting on national interests, the main ones of which are security and survival. To ensure the latter, states are constantly in the pursuit of power, which ultimately leads to the security dilemma. Leading proponents of the classical realist perspective include Hans Morgenthau, E. H. Carr, Reinhold Niebuhr, John H. Herz, Arnold Wolfers, Charles Beard, and Walter Lippman.

Over the last 20 to 30 years, a new form of realism, known as neorealism or structural realism, emerged in response to internal and external debates that challenged key realist assumptions. In particular, neorealism sought to redefine classical realism into a more positivist social science. To that end, neorealist scholars (such as Kenneth Waltz, Stephen Walt, Robert Gilpin, Randall Schweller, John Mearsheimer, Robert Jervis, Joseph Grieco, and Robert J. Art) reformulated realist rhetoric in favor of a systemic approach to international relations. More specifically, neorealism introduced the concept of international structure in order to construct a systems theory explaining what governs relations among states.

Both realism and neorealism are still among the leading schools of thought governing the study of international relations. In addition, they are often invoked by politicians and academicians alike, not only to explain but also to justify state behavior on the international scene. Thus, given their academic and practical importance, the need to fully comprehend the dynamics of realism and neorealism through past scholarly work, as well as their current standing in academia, warrants and justifies the inclusion of this chapter in this *21st Century Political Science: A Reference Handbook*. To do so, this chapter first reviews in a consecutive fashion the evolution and main precepts of both realism and neorealism. Next, it presents key differences between classical realism and neorealism and outlines main neorealist contributions to the study of international relations. Finally, it summarizes several challenges posed to realism and neorealism by proponents of other schools of thought and discusses the current standing of realism and neorealism in the study of international politics.

Evolution and Main Precepts: Realism

Although realism developed as a distinct theory in international relations only around World War II, key realist concepts can be found in much earlier works. Among those are the *History of the Peloponnesian War*, written by the ancient Greek historian Thucydides, who in the 5th century BCE

attributes the real reason for the war to “the growth of Athenian power and Spartan fear of it” (1:23). One century earlier, the ancient Chinese military strategist Sun Tzu also endorsed the role of realpolitik strategizing in states’ behavior in his classic, *The Art of War*. The essence of Thucydides’ realism was recaptured in the writings of the medieval Italian philosopher Niccolò Machiavelli (2008), who, in *The Prince*, emphasized that a politician (i.e., a prince) above all is to be pragmatic (and not idealistic), seeking and using power when necessary to attain practical ends. Other writings promoting realism included *Leviathan*, by the 17th-century English philosopher Thomas Hobbes (2009), wherein he established the concept of anarchy defined as a state of “war of all against all” (1:xiii). Two centuries later, Otto von Bismarck, a Prussian statesman, used and coined the term “balance of power”—a key realist concept depicting a delicate, balancing act among states trying to keep the peace or the status quo in an anarchic environment.

Thus, increasingly, key realist terms such as *national interest*, *security*, *realpolitik*, and *raison d’état* (introduced by Cardinal de Richelieu during the Thirty Years’ War) as well as *balance of power* entered the lexicon of state foreign relations. World War I, however, brought a blow to realism: Woodrow Wilson, the president of a country whose national historical experience had differed substantially from that of the European states, put forward 14 idealistic points aimed at permanently ending the war and establishing peace based on transparency, diplomacy, and honesty. It is also during this postwar era of optimism and pacifism that the study of international relations was first established as an official academic discipline (in 1919) at the University of Wales. Thus, some argue, the climate surrounding its establishment also bestowed a responsibility on academia (and the branch of international relations in particular) to contribute to ending armed conflicts.

However, World War II ended the idealistic optimism that attempted to curtail, if not to end, realist explanations of state behavior. In the climate that followed this war of unprecedented dimensions, E. H. Carr (1939), a historian and former diplomat, wrote *The Twenty Years’ Crisis, 1919–1939*. In it, he called for a return to realism, which he saw as an antidote to the (liberal) utopianism that marked the beginnings of international relations theory in the interwar period. The return and evolution of realism was further fueled by a European immigrant to the United States: Hans J. Morgenthau (1973), whose 1946 publication of *Politics Among Nations* outlined the premises of classical realism.¹ The latter’s theory is based on six principles that Morgenthau delineates in the first chapter of his book, affirming according to classical realism: (1) Politics are governed by objective laws rooted in human nature, which does not change, and it is therefore possible to derive a rational theory reflecting those objective laws; (2) interest, defined in terms of power, is key to understanding the complex tapestry of international politics, but

power, on the other hand, is defined as “anything that establishes and maintains the control of man over man” (p. 13) and covers social, physical, and psychological means; (3) the key concept of interest, defined in terms of power as an objective category, is universally valid but is not endowed with a fixed meaning; (4) prudence, or the “weighing of the consequences of alternative political actions,” not morals, is the ultimate virtue in politics (p. 14); (5) the moral aspirations of a particular nation are not synonymous with the moral laws that govern the universe, but the concept of interest (defined in terms of power), however, prevents moral excess and political folly; and (6) the political sphere is viewed as autonomous, and the only question of importance probing the political life is that based on interests, asking, “How does this policy affect the power of the nation?” (p. 15). Thus, for the political realist, the only relevant standards of thought are the political ones, leaving out moral, economic, legal, and other considerations.

Based on Morgenthau’s six principles as well as on earlier realist concepts, the key precepts of classical realism assert the following: The international system is anarchic; states are its principal, unitary actors; states are rational actors driven by national interests, the ultimate of which are national security and survival; states pursue those interests by amassing power;² power capabilities determine relations among states; and war is a natural, even necessary, occurrence (Carr, 1964; Herz, 1950; Morgenthau, 1973). Realists foresaw that in the pursuit of security through power and arms accumulation, states may face greater insecurity. This, termed as *the security dilemma*, is attributed by realists to the anarchic international environment that prompts all states to fend for themselves (Herz, 1950). In addition, classical realists perceived international relations as governed by objective laws rooted in human nature yet argued that moral principles apply differently to states (Morgenthau, 1973; Niebuhr, 2001). In fact, the inclusion of morality into international relations is seen by realists as problematic and potentially leading to irrational behavior and conflict escalation.

Evolution and Main Precepts: Neorealism

In the first half of the 20th century, classical realism was affirmed as the dominant strand in the classical tradition of international relations theory (with the second being the liberal, or Grotian, tradition, which stressed the impact of concepts such as domestic and international society as well as interdependence and international institutions) (Holsti, 1985). In the early 1970s, current events led many theorists to question traditional concepts of realism. In particular, the widespread opposition to the Vietnam War and the ensuing détente arguably reduced the importance of nuclear competition (Nye, 1988). In addition, the parallel growth of international trade, the spread of transnational

corporations, a decline in U.S. economic predominance, and the oil crisis of 1973 led President Nixon and Secretary of State Kissinger to speak of a potential five-power world while some scholars foresaw the imminent formation of a multipolar international system (Kahn & Bruce-Briggs, 1972). The oil crisis and the vulnerability of the western states during a period of high commodity prices enabled even weak states to extract important resources from the strong. This led even Hans Morgenthau (1974) to observe that “an unprecedented divorce of military and economic power based on the control of raw material” (p. 56) is taking place. In fact, at this period, many political theorists shared concerns that power has undergone a transformation reflected in Brown’s writings:

The forces now ascendant appear to be leaning toward a global society without a dominant structure of cooperation and conflict—a *polyarchy* in which nation states, subnational groups, and transnational special interests and communities would all be vying for the support and loyalty of individuals, and conflicts would have to be resolved primarily on the basis of ad hoc bargaining in a shifting context of power relationship. (cited in Nye, 1988, p. 236)

Thus, realist arguments were challenged not only by current events (which realism seemed ill-equipped to explain) but also by internal debates as well as external, alternative political theories. The latter included the behavioralist movement of the 1960s; the drive toward a more scientific, positivist approach in late 1970s; and the challenge to integrate theory with specific foreign behaviors in the 1980s and 1990s. More specifically, classical realism was particularly criticized for not being scientifically rigorous enough, which, coupled with the rise of alternative political theories, marked a decline in its evolution in the 1960s and 1970s. However, in 1979, in an effort to steer classical realism into a more scientific, positivist direction, Kenneth N. Waltz (1979), through his book *Theory of International Relations*, gave a new vigor to realism by opening a new page in its development, also known as *neorealism*.

Neorealism, also termed *structural realism*, developed from an internal realist debate seeking to address shortcomings of earlier theories of international relations (including reductionism and systemic theories) by developing a more scientific, positivist approach. In Chapter 3 of his book, Waltz (1979) presents the need for developing a systems theory of international politics. He argues that such a theory will describe how the international system works by focusing on its structure and how it affects the interactions among its main units. As such, a systems theory, Waltz writes, would deal with forces at the international, not the national, level and will have both an explanatory and predictive power. In addition, it will be elegant in its ability to apply general concepts to explain and predict continuity and repetition in international politics. Most of all, however, a system theory of international politics will focus on the system per se and not on the attributes or

the sum of the system’s units. Thus, Waltz goes beyond the classical realists’ representation of international politics in terms of states’ characteristics, rationales, and interactions, focusing instead on one level higher—namely, the international system itself, as a unit of analysis. Later, Waltz points out that “the idea that international politics can be thought of as a system with a precisely defined structure is neorealism’s fundamental departure from traditional realism” (p. 27).

In *Neorealism and Its Critics*, Robert Keohane (1986) asserts that the significance of Waltz’s work is not in establishing a new theory but in the systematization of realism or in providing a “more elegant theoretical basis for realism” (Nye, 1988, p. 241). The key accomplishment of structural realism, therefore, consists in the introduction and definition of international structure, which includes three principles. The first is an ordering principle postulating that the international system tends to be either anarchic or hierarchical in nature. In the neorealist view, anarchy is defined as the lack of a central governing body that can arbitrate between states. If the ordering principle of the system transitions from anarchy to hierarchy, the structure of the system will change as well. The second principle is the function of the units (i.e., states). Neorealism perceives states as having similar functions and therefore does not take into account changing state functions when studying the structure of the system. The third principle is the distribution of capabilities across units. By the term *capabilities*, Waltz (1992) refers to the comprehensive and combined material power of a state, including population, economic development, and military force. Based on the distribution of capabilities, neorealists contend that three possible systems may exist: a unipolar, a bipolar, and a multipolar system containing respectively one, two, or more than two great powers. A consensual neorealist argument is that a bipolar system is more stable and thus less prone to war and systemic change than a multipolar system since there are no additional powers with which to form alliances (see also Jervis, 1998).

Classical Realism Versus Neorealism

Classical realism and neorealism differ in four substantive ways, namely by their changing focus on *structure* as a concept, shifting understanding of causality, different interpretations of power, and dissimilar views of the unit level (Waltz, 1992).

First, the idea that international politics can be conceived as a system having a well-defined structure is the main departure point of neorealism from the classical realism. It is the structure, in fact, composed of interacting units with behavioral regularities, that dictates the behavior of its parts. Neorealism argues that the structure of the international system is defined by an ordering principle and by the distribution of capabilities across units (Waltz, 1979, 1992).

In international politics, neorealists affirm, the ordering principle of the international system's structure is anarchy, defined as the absence of a neutral arbiter and higher authority between states. (Yes, ordered anarchy sounds like an oxymoron, but it is indeed key and logical within neorealist theory.) The distribution of capabilities, on the other hand—seen in terms of unipolar, bipolar, or multipolar power configurations—predicts variations in the balance-of-power behavior of states. Thus, Waltz not only provides a systemic theory to explain and predict the behavior of states (the units) but a parsimonious structural theory at that. However, Waltz's theory—and specifically its parsimony—has been critiqued extensively in Keohane's (1986) *Neorealism and Its Critics* and Rosencrance's (1986) *The Rise of the Trading State* (both works are summarized and assessed by Nye, 1988). Waltz (1986), in fact, offers his counterarguments in defense of neorealism in his "Reflections on *Theory of International Politics: A Response to My Critics*."

One of the criticisms of Waltz's neorealist systems theory is its vaguely defined concept of *international structure*. Keohane (1986) has argued that neorealist theory is too adaptable, for it "can be modified progressively to attain closer correspondence with reality" (p. 191). Barry Buzan (1988), on the other hand, questions whether the structural logic of neorealism captures "the main features of the international political system" (p. 35). He writes, "The criticisms of Ruggie, Keohane, and others suggest that it does not, because their concerns with factors such as dynamic density, information richness, communication facilities, and such like do not obviously fit into Waltz's ostensibly 'systemic' theory" (cited in Waltz, 1992, p. 30).

Waltz (1992) responds to such criticisms by affirming that concepts such as *dynamic density*, *information richness*, and *communication facilities* are not and cannot be elements of a theory, in general, or of his systemic, structural, neorealist theory, in particular. Rather, those are conditions that develop within nations, across nations, or both, that may disrupt and even transform respective societies or cross national relations. Yet such concepts do not and cannot define neorealist theory. In general, Waltz argues, a theory, if it is a good one, would help to understand and explain such concepts' significance and effects within and on the system. However, he affirms, a theory cannot fit the facts or concepts that it seeks to explain. In other words, a theory can be written only by omitting most matters that are of practical interest. To criticize the neorealist systems theory based on its omissions is to misconstrue the essence and purpose of a theory.

Second, classical realism and neorealism differ in their views of causality in international politics—in other words, what causes the observed outcomes in relations among states. For classical realists, the international world is one of interacting states, and causes run in one direction: from interacting states to the outcomes their acts and interactions produce. Neorealists, on the other hand, adopt a

more deductive approach by distinguishing between structural and unit-level causes and effects in order to study interacting states. As such, neorealism contends that international politics can be understood only if the effect of the international structure is added to the classical realism's unit-level analysis. Neorealist theory adopts, therefore, a two-directional causality running from interacting units to produced outcomes, on one hand, and from the structural level to interacting units, on the other. Neorealists contend that the inclusion of both unit-level and structure-level causes allow for neorealist theory to "cope with both the changes and the continuities that occur in a system" (Waltz, 1992, p. 34).

A third dissimilarity between classical realism and neorealism lies in their different interpretations of power. Morgenthau (1973) recognized that scarce resources and the lack of neutral arbiter would lead to a struggle for power among competitors, which speaks of the latter's nature. Classical realists saw states' desire for and the pursuit of power as ultimately rooted in the nature of man. In addition, even when one has accumulated much power, Morgenthau asserts that more is still needed:

Since the desire to attain a maximum of power is universal, all nations must always be afraid that their own miscalculations and the power increases of other nations might add up to inferiority for themselves which they must at all costs try to avoid. (p. 208)

Thus, in seeking to explain state behavior internationally, Morgenthau (1973) conceives it as that of the rational statesman acting with prudence and striving to accumulate more and more power. Power is the ultimate goal and an end in itself. When nations act outside of power considerations, Morgenthau asserts that those actions are not "of a political nature" (p. 27). Thus, the claim that "the desire to attain a maximum of power is universal" (p. 27) is among Morgenthau's objective laws that are ultimately rooted in human nature. Neorealists, on the other hand, rather than seeing power as an end in itself, perceive it as a potentially useful means of which states may have either too little or too much (Mearsheimer, 2001; Waltz, 1979, 1992). Further, neorealists use the concept of power as one of the central characteristics of the system and define it as the combined capabilities of a state.

The fourth main difference between realism and neorealism is their treatment of the unit level. Since realists perceive anarchy to be a general condition (rather than a distinct structure) of the international system, they see interacting units acting in response to their anarchic environment. The different outcomes are attributed to differences within the units, such as dissimilar forms of government, rulers, and types of ideology. Neorealists, on the other hand, affirm that "states are made functionally similar by the constraints of the structure, with the principal differences among them defined according to capabilities" (Waltz, 1992, p. 36). In other words, neorealism

argues that it is structure that mediates the outcomes that states produce and not exclusively the internal state characteristics. Thus, neorealists concentrate on a previously unanswered question in the study of international politics—namely, “How can the structure of an international-political system be distinguished from its interacting parts?” (p. 37). Neorealists perceive states like units, with each state being like all other states in its political autonomy. Thus, while realists concentrate on the heterogeneity of states to explain differences in external behavior, neorealists also add the effects that structure has on state behavior and outcomes.

In summary, although neorealism, pioneered by Waltz, concurs with classical realism that states are the principal actors, it argues that the conditions of the system as a whole influence state behavior and not only state interactions per se. The international system is seen as a structure that causes and defines relations among states. Neorealists reject reductionist theories and, instead, view states as so-called black boxes, thereby dismissing their domestic characteristics as having an exclusive role in states’ foreign behavior (Mearsheimer, 2006; Waltz, 1979). In addition, neorealists, similarly to the classical realists, assert that conflict and war are unavoidable and permanent occurrences in relations among states. However, while classical realists view war as rooted in human nature, neorealists argue that it is the anarchic structure of the international system characterized by the absence of a neutral authoritative body enforcing rules and agreements between states that causes and perpetrates interstate wars.

Beyond Waltz

Thanks to several key contributions, neorealism evolved beyond Waltz’s (1979) pioneering *Theory of International Politics*. In *The Origins of Alliances*, for instance, the neorealist scholar Stephen Walt (1990) introduces the “balance of threat” thesis by expanding on the realist concept of the balance of power. He argues that since states exist in an anarchical self-help system, they tend to balance not against a rising power but rather against a perceived threat. To strengthen his arguments, Walt examines 36 bilateral and multilateral alliance formations and 86 national decisions in the Middle East from 1955 to 1979. Following the examination of these alliances, Walt affirms that his balance of threat thesis gives a better understanding of alliance formation in the Middle East than that provided by variables such as ideology, foreign aid, and political penetration. In addition, he observes that balancing rather than bandwagoning is the aim of alliances formed in the Middle East for the examined period.

Also associated with structural realism is Randall Schweller, who developed the “balance of interests” thesis reviewing Waltz’s balance of power and Walt’s balance of threat concepts. Schweller (1993, 1998) expands his arguments on this subject in two main works—namely,

“Tripolarity and the Second World War” in *International Studies Quarterly*, and *Deadly Imbalances: Tripolarity and Hitler’s Strategy of World Conquest*. In addition, contrary to Waltz, Schweller (1994, 1996) differentiates between states in terms of pursued goals. As such, he classifies them in two categories: (1) status quo and security maximizers or (2) revisionist and power maximizers. Schweller demonstrates that the end goals motivate states as well, and they may align not only to balance (against a loss) but also to gain. However, since he is willing to also consider nonstructural explanations of state behavior (Schweller, 1994; Schweller & Press, 1997), he may more accurately be defined as a neoclassical realist (see subsequent discussion).

In addition to this internal division as to the (exclusive) role played by the international structure on relations among states, neorealists are further divided as to how state security must be ensured. To that end, the two main neorealist branches are offensive and defensive realism: While the former defends aggressive state behavior to protect national security, the latter upholds preventive measures. In addition, John Mearsheimer, a leading proponent of offensive realism, blames security competition among states not on human nature (an argument advanced by the classical realism of Hans Morgenthau) but on the anarchy of the international system. In contrast to his defensive neorealist colleague, Kenneth Waltz, Mearsheimer questions the permanency of the status quo, arguing instead that states are never satisfied with power but ultimately seek hegemony to ensure their security. This argument is clearly expressed in Mearsheimer’s (2001) work *The Tragedy of Great Power Politics*, asserting the following:

Given the difficulty of determining how much power is enough for today and tomorrow, great powers recognize that the best way to ensure their security is to achieve hegemony now, thus eliminating any possibility of a challenge by another great power. Only a misguided state would pass up an opportunity to become hegemon in the system because it thought it already had sufficient power to survive. (p. 35)

Thus, offensive realists argue that, in international politics, aggression is a natural outcome that is generated by rational power maximizers, which are all states. Also, in contrast to Schweller, Mearsheimer stays at the structural level, arguing that it is the structure of the international system that causes states to behave offensively rather than perceived gains or losses (Mearsheimer, 2006).

Challenges to (Neo)Realism

Some have suggested that, in modern international relations, realism has become obsolete (Lebow, 1994; Legro & Moravcsik, 1999; Russett, 1993; Vasquez, 1997). These scholars assert that realism’s concepts of anarchy, self-help, and the balance of power are little, if at all, applicable

to the current state of international environment marked by the spread of democracy, the growth of interdependence, and the influence of international institutions. Waltz, specifically, has been long criticized for not incorporating into his system theory unit-level processes that may generate system-level change (a long-standing criticism in this aspect is that neorealism is poorly equipped to explain system change).

However, realist and neorealist scholars argue that although changes in the system have indeed occurred at the unit level, changes of the system have not (Waltz, 2000). Unit-level, within-system changes, such as changes in technology, transportation, communication, and war fighting, occur all the time, and they do affect how states interact. Nuclear weapons, for instance, have decisively altered how states provide for their security, yet nuclear weapons have not changed the anarchic structure of the international system. And only when changes of the system have taken place can one begin to consider whether realism has run its course. However, if the international system's nature has remained unaltered, realist concepts still apply.

Yet some have raised the question whether realist premises really explain the most important phenomena in international relations. As such, in present-day international relations, key realist assumptions have been challenged through several modern phenomena. First, the argument advanced by (neo)realists that long-lasting peace is unattainable because of the anarchic structure of the international system is challenged by the democratic peace theory. The latter argues that realist predictions do not necessarily apply to democratic states and their relations to other democracies (Owen, 1994). Realists have critiqued this claim by pointing out several of the theory's causal inconsistencies as well as weaknesses of its definitions of *war* and *democracy* that can and must be adjusted in order to demonstrate the key claim that democracies do not fight other democracies (Rosato, 2003; Waltz, 2000). Yet the latter statement has been an observed fact that both realists as well as proponents of other international relations schools have yet to explain.

The constructivist school has also challenged realism and neorealism. Through several influential articles, Alexander Wendt (1987, 1992, 1995, 1999) provided perhaps one of the most ardent constructivist critiques of neorealism. In his articles and subsequent *Social Theory of International Politics*, Wendt challenges the core neorealist premise that anarchy forces states into recurrent security competition. Wendt argues that anarchy and power do not define whether a system is conflictual or peaceful; instead, it is shared culture formed through discursive social practices that define it. Since each state's conception of self (including its interest and identity) is shaped through internal processes, states can reshape the structure through new gestures and reconstituted interests and identities and move toward more other-regarding and peaceful means and ends. Thus, Wendt (1992) argues: "Anarchy is

what states make of it," or in other words, states are not forced by their anarchic environment to endlessly amass power and fall into unavoidable conflicts. Instead, focusing on process rather than structure, Wendt argues that states' identities and interests can be collectively transformed by many factors at the individual, domestic, international, and transnational levels and, as such, are an important dependent variable that state-centered theories need to consider.

Despite Wendt's important insights into realist–neorealist polemic, realists argue that his critique has several flaws. First, it does not address a key neorealist concept—namely, that of uncertainty. For structural realists, states' uncertainty as to the present and future intentions of others is what places the concept of (relative) power as a central causal variable. The main issue, neorealists affirm, as to why systems sometimes fall into conflict is not the differently constructed interests, as Wendt affirms, but the uncertainty about other states' interests and motives that leads to the pursuit of security and, potentially, to conflict (Copeland, 1999–2000; Fearon, 1995). In addition, Wendt does not address incentives that actors may have to deceive one another, which may further exacerbate a situation (Copeland, 2000).

Another challenge to realism is posed by the development of critical theories (including feminist approaches) that question the reasoning behind the formulation of key realist concepts (Tickner, 2005). Further, nowadays, with the advance in both theory and practice of such concepts as complex interdependence, globalization, global justice, transnational, and nonstate entities, realist assumptions accounting solely for the role of states as unitary, sovereign actors are challenged to accommodate new political realities. For instance, the presence of federal states wherein sovereignty is constitutionally split between the national and local governments (with citizens having political obligations to both authorities) challenges realism to account for the role of sovereign subunits that may also participate in the central decision making of a state.

Finally, realism and neorealism are positivist approaches focusing on state security and the pursuit of power above all. Realist insistence to study the impact of material forces (such as the size of military forces, and the balance of powers) through a replication of scientific methods has been increasingly challenged by post-positivist epistemology rejecting the idea that the social world can be studied in an objective, value-free way. Postpositivist theories challenging central realist tenets, such as rational choice, argue that scientific methods are inapplicable to the world of social interactions, making it impossible to build a science of international relations (Bernstein, Lebow, Stein, & Weber, 2000). Those arguments, hitting at the very core of realism—namely, its methodology—have grown in preponderance and in adherents over the last two decades. In particular, studies examining the role of psychological factors in international relations have grown in number and importance over the

last few years. Evaluating these factors has come from the rejection of the realist representation of states as black boxes and the argument that there may be other influences on foreign policy decisions that come from within the state. Some of those internal influences are pointed to come from psychological factors, such as the concept of groupthink, the propensity of policymakers to decide based on (mis)perceptions and analogies, or both (Jervis, 1968; Stein, 2002).

In response to some of the above criticisms, neoclassical realism evolved as the third wave of realism (following neorealism). Neoclassical realism does not reject Waltz's neorealist assumptions but rather refines them in order to offer explanations about specific states' behaviors. In particular, neoclassical realists examine how the distribution of power in the international system, together with states' domestic incentives and perceptions of that system, shape their foreign policy. Thus, neoclassical realists replace the neorealist claim that the possibility of conflict shapes the actions of states with assessment of probability, referring to how likely the worst-case scenario is to occur. Similarly, to the neorealist view that short-term gains matter, neoclassical realists add a consideration of long-term state objectives. Finally, to the neorealists' emphasis of state military preparedness, neoclassical realists add economic gains that may outweigh potential security losses (Brooks, 1997).

Although neoclassical realism has been most useful in theories of foreign policy, Schweller (1994, 2003) argues that it can be used to explain other political outcomes as well. Thus, neoclassical realism has been a positive development of realist theory, for it combines the theoretical rigor of Waltz while allowing for a robust method to test theories on specific case studies of states' foreign behavior. Other prominent neoclassical realists include Fareed Zakaria, Thomas J. Christensen, and William Wohlforth.

Conclusion

Realists today disagree on many internal issues, yet they are united in what they agree on—particularly when faced with alternative international relations theories. Namely, they see international relations as relating to objective conditions; they reject ideological, psychological, and normative considerations to explain relations among states that they perceived as defined by the anarchic environment surrounding states; and they view military capabilities and power accrual as paramount to both states' positioning within the international system and their survival and security.

In 1997, Michael Doyle (1997) expressed what is considered the conventional wisdom that “realism is our dominant theory. Most international relations scholars are either self identified or readily identifiable realists” (p. 41). However, a study conducted by Maliniak, Oakes, Peterson, and Tierney (2007) of the current state of the

international relations discipline (from 1980 to 2006) reveals rather interesting results as to the current standing of realism in the academic community—namely, that “the share of published work that fits squarely in the realist tradition” (p.11) is relatively small. This observation goes against a widespread belief among scholars that realism is the most prominent and popular approach in international relations. In fact, despite the fact that realism is highlighted in academia, Maliniak et al. demonstrate that for the past 27 years, “Realism has never been the most popular paradigm for journal authors, and in 1999 it fell to third behind constructivism” (p. 12). In particular, over the past 10 years, realist arguments seem to have been declining, as observed through published articles in the top 12 journals of the international relations discipline. However, Maliniak et al. demonstrate that even though the number of realist articles has declined, nonrealists continue to consider realist approaches seriously, against which they test their own models of state behavior. Thus, Maliniak et al. argue that “realism still looms large in the minds and research designs of non-realist IR scholars” (p. 12). In sum, directly or indirectly, realism is still much present in current international relations debates, research, and teaching.

Notes

1. Michael C. Williams's edited volume *Realism Reconsidered* (2008) takes a critical look on Morgenthau's scholarship on political realism from a historic, intellectual, and theoretical viewpoint. As such, the chapters written by Chris Brown and William E. Scheuerman discuss Morgenthau's exchanges with the controversial German jurist Carl Schmitt; Anthony F. Lang Jr. examines Aristotle's influence on Morgenthau; Richard Little analyzes the balance of power in “Politics Among Nations”; Michael Cox writes on the purported failure of realism to predict the end of the cold war; and Michael Williams examines the juxtaposition of neoconservatism and realism. Generally, all 10 authors note that Morgenthau's realism has been frequently misrepresented as an amoral, and even immoral, theoretical approach, arguing that his writings are much more nuanced and complex than his critics assert. In fact, the authors argue that Morgenthau's rationales resemble in several aspects some advanced by constructivist international relations theories.

2. Since power is a central tenet in realism, the latter is also known as power politics.

References and Further Readings

- Bernstein, S., Lebow, R. N., Stein, J. G., & Weber, S. (2000). God gave physics the easy problems: Adapting social science to an unpredictable world. *European Journal of International Relations*, 6(1), 43–76.
- Brooks, S. G. (1997). Dueling realisms. *International Organization*, 51(3), 445–477.
- Buzan, B. (1988). *Systems, structures, and units: Reconstructing Waltz's theory of international politics*. Unpublished manuscript.

- Carr, E. H. (1964). *The twenty years' crisis, 1919-1939: An introduction to the study of international relations*. New York: Harper & Row.
- Copeland, D. C. (1999-2000). Trade expectations and the outbreak of peace: Détente 1970-74 and the end of the cold war 1985-91. *Security Studies*, 9(1-2), 15-58.
- Copeland, D. C. (2000). The constructivist challenge to structural realism. *International Security*, 25(2), 187-212.
- Doyle, M. (1997). *Ways of war and peace: Realism, liberalism, and socialism*. London: W. W. Norton.
- Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, 49(3), 379-414.
- Herz, J. (1950). Idealist internationalism and the security dilemma. *World Politics*, 2(2), 157-180.
- Hobbes, T. (2009). *Leviathan*. Oxford, UK: Oxford University Press.
- Holsti, K. J. (1985). *The dividing discipline: Hegemony and diversity in international theory*. Boston: Allen & Unwin.
- Jervis, R. (1968). Hypotheses on misperception. *World Politics*, 20(3), 454-479.
- Jervis, R. (1998). Realism and the study of world politics. *International Organization*, 52(4), 971-991.
- Kahn, H., & Bruce Briggs, B. (1972). *Things to come*. New York: Macmillan.
- Keohane, R. (Ed.). (1986). *Neorealism and its critics*. New York: Columbia University Press.
- Lebow, R. N. (1994). The long peace, the end of the cold war, and the failure of realism. *International Organization*, 48(2), 249-277.
- Legro, J. W., & Moravcsik, A. (1999). Is anybody still a realist? *International Security*, 24(2), 5-55.
- Machiavelli, N. (2008). *The prince*. Miami, FL: BN Publishing.
- Maliniak, D., Oakes, A., Peterson, S., & Tierney, M. J. (2007, August-September). *The international relations discipline, 1980-2006*. Paper presented at the annual meeting of the American Political Science Association, Chicago.
- Mearsheimer, J. (2001). *The tragedy of great power politics*. New York: W. W. Norton.
- Mearsheimer, J. (2006, March). Conversations in international relations: Interview with John J. Mearsheimer (part I). *International Relations*, 20(1), 105-123.
- Morgenthau, H. J. (1973). *Politics among nations* (5th ed.). New York: Knopf.
- Morgenthau, H. J. (1974). The new diplomacy of movement. *Encounter*, 43, 52-57.
- Niebuhr, R. (2001). *Moral man and immoral society: A study of ethics and politics*. Louisville, KY: Westminster John Knox Press.
- Nye, J. S. J. (1988). Neorealism and neoliberalism. *World Politics*, 40(2), 235-251.
- Owen, J. M. (1994). How liberalism produces democratic peace. *International Security*, 19(2), 87-125.
- Rosato, S. (2003). The flawed logic of democratic peace theory. *American Political Science Review*, 97(4), 595-602.
- Rosenance, R. (1986). *The rise of the trading state*. New York: Basic Books.
- Russett, B. (1993). *Grasping the democratic peace: Principles for a post cold war peace*. Princeton, NJ: Princeton University Press.
- Schweller, R. L. (1993). Tripolarity and the Second World War. *International Studies Quarterly*, 37(1), 73-103.
- Schweller, R. L. (1994). Bandwagoning for profit: Bringing the revisionist state back in. *International Security*, 19(1), 72-107.
- Schweller, R. L. (1996, Spring). Neorealism's status quo bias: What security dilemma? *Security Studies*, 5(3), 90-121.
- Schweller, R. L. (1998). *Deadly imbalances: Tripolarity and Hitler's strategy of world conquest*. New York: Columbia University Press.
- Schweller, R. L. (2003). The progressiveness of neoclassical realism. In C. Elman & M. F. Elman (Eds.), *Progress in international relations theory* (pp. 311-347). Cambridge: MIT Press.
- Schweller, R. L., & Press, D. (1997). A tale of two realisms: Expanding the institutions debate. *Mershon International Studies Review*, 41(2), 1-32.
- Stein, J. G. (2002). Psychological explanations of international conflict. In W. Carlsnaes, T. Risse, & B. Simmons (Eds.), *Handbook of international relations* (pp. 292-308). Thousand Oaks, CA: Sage.
- Tickner, J. A. (2005). What is your research program? Some feminist answers to IR's methodological questions. *International Studies Quarterly*, 49(1), 1-21.
- Vasquez, J. (1997). The realist paradigm and degenerative vs. progressive research programs: An appraisal of neotraditional research on Waltz's balancing proposition. *American Political Science Review*, 91(4), 899-912.
- Walt, S. (1990). *The origins of alliances*. Ithaca, NY: Cornell University Press.
- Waltz, K. (1979). *Theory of international politics*. New York: McGraw Hill.
- Waltz, K. (1986). Reflections on theory of international politics: A response to my critics. In R. Keohane (Ed.), *Neorealism and its critics* (pp. 322-346). New York: Columbia University Press.
- Waltz, K. (1992). Realist thought and neorealist theory. In R. L. Rothstein & W. T. R. Fox (Eds.), *The evolution of theory in international relations* (pp. 21-38). Columbia: University of South Carolina Press.
- Waltz, K. (2000). Structural realism after the cold war. *International Security*, 25(1), 5-41.
- Wendt, A. (1987). The agent structure problem in international relations theory. *International Organization*, 41(3), 335-370.
- Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 42(2), 391-425.
- Wendt, A. (1995). Constructing international politics. *International Security*, 20(1), 71-81.
- Wendt, A. (1999). *Social theory of international politics*. Cambridge, UK: Cambridge University Press.
- Williams, M. (2008). *Realism reconsidered: The legacy of Hans Morgenthau in international relations*. Oxford, UK: Oxford University Press.

38

IDEALISM AND LIBERALISM

ERIC A. HEINZE

University of Oklahoma

BRANDY J. JOLLIFF

University of Colorado at Boulder

Most textbooks on international relations (IR) characterize liberalism as one of the main theoretical schools of the IR field—typically alongside realism and perhaps some other less mainstream approaches like international society, Marxism, constructivism, or feminism. As such, liberalism is commonly considered to be the main competing theoretical approach to the dominant IR theory of realism. The frequent comparisons made between realism and liberalism in the IR literature typically entail realism advancing a pessimistic view of human nature, versus the more optimistic view espoused by liberalism. Realists therefore see conflict as the norm in international affairs, while liberals are more hopeful about the prospects for peace and international cooperation. Realists seek to explain international politics by examining state-to-state relations within an anarchical system of mutual distrust and suspicion, while liberals consider other international actors, as well as actors and institutions within the state, as the underlying causes of a more interdependent and law-governed world.

This broad understanding of liberalism represents the approach as it has developed throughout the post–World War II era. Although contemporary liberal theory can be divided into different strands, which this chapter discusses in a following section, the notion of idealism as it pertains

to IR is a slightly different and older idea that played an important role in the evolution of what is now recognized as contemporary liberal IR theory. Idealism—sometimes referred to as utopianism—was a popular approach to analyzing international politics in the period immediately following World War I. It was identified as a theoretical tradition of IR largely in hindsight, with the various attempts by realists at discrediting its central tenets, which were caricatured as utopian or idealistic (see Carr, 2001; Morgenthau, 1993). Although it is true that what is recognized as liberal IR theory has intellectual roots in the idealist tradition of the interwar period, both idealism and contemporary liberalism have their origins in European Enlightenment political thought. This chapter thus traces the intellectual origins of contemporary liberal IR theory to the modern liberal philosophers who theorized about the state. It then describes how liberal theories of the state came to be applied to international politics, subsequently caricatured as idealist, and how the liberal-idealist approaches informed attempts at creating international institutions and organizations. The chapter then discusses how liberal theory enjoyed a revival after the end of the cold war and outlines the different strands of liberal theory that have emerged since World War II. This chapter concludes with a discussion of the main international issues and challenges that confront contemporary liberal IR theory.

The Intellectual Roots of Idealism and Liberalism

Liberalism and Enlightenment Thought

The driving force behind liberalism as a political theory of the state is the centrality of individual liberty. The liberal ideal entails a limited or conditional government, whose legitimacy is derived from the consent of the governed, over whom rulers may not exercise coercion except through means established by law. Liberalism thus espouses a concept of the state whose job it is to remove obstacles to freedom and protect individuals from even majoritarian oppression. To prevent governments from exceeding these limits, of course, requires the familiar array of institutional constraints, checks and balances, and individual rights that underlie the constitutional arrangements of nearly every liberal-democratic polity that exists today.

The English philosopher John Locke (1800) advanced this idea of a limited government using a common state-of-nature argument, whereby all individuals in the state of nature—the prepolitical existence of humans before we lived under the authority of government—had “natural rights” to life, liberty, and property. It is through the human capacity for reason that individuals are aware of such law, though without government to enforce it, transgressors of natural law may be pursued and punished by any person who lives according to the laws of nature, not just by those whose rights were unjustly deprived. The problem, of course, is that such individuals are unlikely to be fair and impartial when punishing transgressors, which is precisely why Locke argued that rational individuals would establish civil government, though one that would preserve and protect the freedoms that individuals had in the state of nature. The idea of a liberal state as it emerges from this Lockean analysis is therefore characterized by political freedom, democracy, constitutionally protected rights, as well as private property.

Many subsequent theorists in the liberal tradition took up Locke’s arguments about the proper structure of commonwealths and began applying them to relations among commonwealths. Modern legal theorists such as Emmerich de Vattel (1863) have been associated with a distinctively Lockean analysis of international relations in that states have no government to rule over them or enforce their rights but are governed by a universal natural law (which Vattel termed the *necessary law of nations*) that is binding on all states and obligates them to respect the rights of one another. Indeed, Vattel uses this general framework as a way to conceive of what we would today refer to as international law and collective security, both of which are widely recognized as liberal prescriptions for international relations.

Immanuel Kant was another important figure in the application of liberal theory to international relations and

is commonly cited as one of the founding fathers of idealism (Hutchings, 1999). Building from the Lockean liberal ideas of individual liberty and popular sovereignty, as well as the Enlightenment credo of human progress and perfectibility, Kant is best known for arguing that states with republican constitutions (i.e., liberal, democratic states) are inherently more peaceful and will thus design international laws to regulate interstate behavior and to promote the conditions for peace. The fundamentally Kantian insight that the domestic politics and institutions of states are critical factors in explaining their international behavior is perhaps the defining feature of liberal IR theory and is the central component of what is widely known as the democratic peace theory (DPT). Kant also argues for the creation of an international federation of democratic, peaceful states that will expand its membership over time and make the world more peaceful. Kant is not calling for a world government, but rather a sort of loose union of states that maintains itself, prevents war, and steadily increases its membership (Kant, 1991).

Locke’s conception of private property was likewise an important starting point for much theorizing on the ideas that free and open societies should have an open marketplace. This is not only because market capitalism was thought to best promote overall welfare by efficiently allocating scarce resources within society, but also because of the supposed pacifying effects that this has internationally. According to liberal thinkers such as Jeremy Bentham, John Stuart Mill, and Adam Smith, free and open trade among nations has a general harmonizing effect since it is mutually beneficial and contributes to the happiness of one another’s society (see Howard, 1978). The basic idea was twofold. First, since many wars were fought by states as a means to enrich themselves yet these wars still proved to be costly and did not benefit the society as a whole, free trade would be a more peaceful and efficient means of achieving wealth, which is a common interest of all states. As a corollary, the free movement of commodities, capital, and labor across borders would break down divisions between states. This would open up lines of communication between them to reduce uncertainty, binding countries together using the common tie of economic interest (Ricardo, 1911). Thus, what Michael Doyle (1997) refers to as “commercial pacifism” is simply the idea that market societies are fundamentally against wars.

What emerges from this discussion of international liberalism as it evolved from Lockean liberalism throughout the Enlightenment is a set of ideas about international relations that include (a) a strong preference for a law-governed society of states, (b) cooperation in international organizations to collectively enforce this law, (c) the spread of democracy and liberal values (therefore bringing about peace), and (d) the pursuit of free trade to enhance global prosperity and help bring about peace. Thus, by the late 19th and early 20th centuries, liberal Enlightenment thought had produced the broad contours of what would

become known as idealism and set the stage for the emergence of IR as an academic discipline.

Idealism and the Twenty Years' Crisis

The calamity of World War I and the horrific human toll it brought about led to new efforts to try and understand, prevent, and ultimately eliminate war. As a result, the academic discipline of IR was born. For liberal thinkers of the time, the war was largely a result of the egoistic and reckless miscalculations of autocratic rulers in heavily militarized countries, as well as the outdated system of alliances based on a balance of power that had dominated Europe for centuries (see Jackson & Sorensen, 2007). Since liberal thinkers had some clear ideas and strong beliefs on how to avoid such disastrous wars in the future, the emerging discipline of IR was highly influenced by these liberal principles and was guided by a desire to replace the malfunctioning European balance of power with a system of international law and collective security, as well as to reform the structure of autocratic governments in order to make them more peaceful.

Prominent among this group of liberal intellectuals was British writer Norman Angell (1913), whose book *The Great Illusion* argued that war was no longer a profitable and useful tool for the conduct of state foreign policy. Angell argued that wars of conquest between industrialized states had become futile and that the best solution to aggression was “third party judgment” within a collective system (Miller, 1995). For Angell, states’ single-minded pursuit of their own security in a condition of anarchy (i.e., the absence of a world government) led to war; thus, security needed to be provided internationally. After the Great War, he became an ardent supporter of the League of Nations, suggesting that “the military power of the world should be so pooled by international agreement for supporting a common rule of life for the nations as in fact to make it the police power of civilization” (cited in Miller, 1995, p. 112).

Yet it is perhaps the architect of the League of Nations—U.S. President Woodrow Wilson—who is most commonly associated with interwar idealism. President Wilson entered the United States into World War I on a decidedly liberal platform: to make the world safe for democracy. Wilson was highly critical of the European balance-of-power system and saw it as his mission to bring liberal democratic values to the rest of the world. Wilson’s Fourteen Points contained his vision for the new liberal foundation of international politics, which emphasized, *inter alia*, the promotion of democracy and self-determination based on the conviction that democracies do not go to war against each other. Another important principle contained in Wilson’s vision was the creation of an international organization based on a set of common rules in international law that would replace the unstable balance-of-power system that he argued had failed to prevent the war.

The League of Nations was therefore created to promote peaceful cooperation among states based on the idea that there should be reason-based substitutes for war. Although the realists were content to allow the dangerous game of power politics to occur unrestrained based on an unstable balance of power, Wilson’s view was that the warlike impulses of states, statesmen, and other instruments of conflict could be controlled by an intelligently designed international institution. This notion of Wilsonian idealism was thus based on the liberal view that when rational human beings apply reason to international problems, they can establish institutions that can improve the human condition (Jackson & Sorensen, 2007).

Highly influenced by Wilsonian idealism, IR scholarship during the interwar period consisted mainly of forward-looking liberal conceptions of world federations, blueprints for a more perfect League of Nations, and the development of new international institutions and legal codes for interstate behavior, all amid a strong normative desire for the avoidance of great-power war (Wilson, 1995). Yet as we know, the League of Nations was doomed to failure, and the ideas being championed by the likes of Angell and Wilson came under intense criticism. The League was helpless against the onslaught of the Great Depression and the protectionist policies that ensued, as well as the expansionist policies of Germany, Japan, and Italy. Perhaps the best known critique of the interwar idealists is that of E. H. Carr’s *Twenty Years’ Crisis* (2001), which is most famous for its attempt to debunk the pretensions of the liberal thinking that dominated the international relations discourse during the twenty years’ crisis, between 1919 and 1939. Carr argued that liberal thinkers had fundamentally misread history and therefore misunderstood the nature of international relations (Knutsen, 1997). Although the idealists believed that international relations could be based on a harmony of interests among different states, Carr argued that this was wishful thinking (hence utopian) and that we should assume that there are conflicts of interests among states. In short, Carr accused the liberals of being too preoccupied with what international relations ought to resemble rather than what it actually resembled and for overemphasizing the role of international law and morality and underestimated the role of power (Carr, 2001; Wilson, 1995). This framework, which posed a dialectic between utopia and reality, would be highly influential in the development by later realists of a more scientific, fact-based way of studying IR that emerged in the 1950s with the behavioral revolution in the social sciences (Waltz, 1979). With the spread of autocratic and militaristic states and the failure of the League to prevent the outbreak of World War II, the liberal assumptions underlying Wilsonian idealism fell out of favor among IR scholars, and the field soon became dominated by realist thinking, with its pessimistic view of human nature and emphasis on international relations as a conflictual struggle for power within an anarchical system. Yet liberal thinking

remained an influential part of IR theory and would soon reemerge as an important source of scholarship as it was refined in light of the realist challenge to its foundational principles.

Contemporary Liberal IR Theory

Liberalism in the Postwar Era

The bipolar structure of the cold war period put considerable stress on liberal theory's ability to explain international politics, since realism arguably offered more explanatory power in the context of an anarchical system dominated by two powerful hegemonies mired in a security dilemma (see Waltz, 1979). Yet as international actors emerged from World War II and were forced to confront pressing issues about the future international political and economic order, liberal principles continued to play a prominent role. The postwar order was fundamentally organized as a rule-based international order, wherein international cooperation was encouraged as a means to ensure peace, economic prosperity, and human rights. Such was the rhetoric of the founding treaties of many postwar international organizations (IOs), such as the United Nations (UN), European Community (EC), and the Bretton Woods institutions. Although not solving the world's problems, the interstate cooperation that these organizations encouraged gave liberal IR scholars renewed optimism about the role international institutions could and should play in world politics and provided a whole new set of organizations, institutions, regimes, processes, and interactions that became the subject of investigation by liberalist IR scholars.

Despite the emergence of several new international organizations during the postwar era, the international security environment was dominated by cold war power politics. Yet at least in the West, the Bretton Woods institutions and the UN offered glimmers of hope to those still attempting international cooperation in a threatening, hard, power-dominated bipolar system. These key institutions were created to govern monetary relations among the world's states, to encourage free trade among them, and ultimately to facilitate the spread of free market economics.

As IR scholars began considering the power realities of the postwar period—particularly the hostility between the United States and the Soviet Union—the idealism that dominated interwar thinking gave way to realism, particularly among U.S. academics, which was further fueled by the rise of behavioralism in political science. The rise of behavioralism in the social sciences was essentially a call for more rigorous methodologies that applied stricter, more scientific reasoning in IR scholarship that was to be less normative and ideologically driven and more interested in observable facts, measurable data, and the finding

of “law-like” behavioral patterns (Knutsen, 1997). Although the supposedly more objective and dispassionate realism was perhaps a better fit to such a method for social-science scholarship, new formulations of both realism and liberalism emerged in an attempt to answer the call of the behavioralists for more methodological rigor.

Inspired by the scientific ambitions of behavioralism, Kenneth Waltz (1979) developed a new form of realism—dubbed *neorealism*—that focused on the structure of the international system comprised of unitary states, wherein he attempted to achieve lawlike statements about international politics that could achieve scientific validity. For Waltz, the anarchical structure of the system leads rational states to be power seeking and inherently distrustful of other states, thus leading to the fundamentally conflictual character of international politics. Liberal theorists, such as Robert Keohane and Joseph Nye, took a slightly different approach in their attempts to answer the behavioralist challenge. This work was largely based on the early functionalist research of those scholars studying European integration in the 1950s, who studied how cooperation in one issue area can “spillover” to allow for cooperation in other areas (Deutsch, 1957; Haas, 1958; Mitrany, 1966). This new brand of liberalism—or neoliberalism—sought to explain the various instances of cooperation among democratic states by reference to the idea of *complex interdependence*, which includes the various forms of connection between states in addition to the political relations of their governments, such as transnational links between businesses (Keohane & Nye, 1971, 1977). This leads to an absence of hierarchy among issues—that is, a condition where military security is not necessarily states' top priority. Thus, in contrast to the neorealist vision of international politics, the neoliberals argued that there are other important actors in international relations that contribute to interdependence among states, which leads to less conflict among them.

In such an interdependent world, openings developed for international institutions and IOs to become influential actors that facilitate cooperation through information exchanges and the provision of arenas for resolving disputes. This became the basis for another wave of neoliberal IR scholarship that focused on the role that international organizations and regimes played in state behavior (Krasner, 1983). It was Robert Keohane's (1984) *After Hegemony* that was perhaps the most influential publication on these general themes. Seeking to address the neorealist critique of neoliberalism head-on, Keohane adopts many of the foundational assumptions of neorealism. Whereas the neorealists argue that this rationality leads to conflict, Keohane demonstrates that it can lead to cooperation and the establishment of institutions. Building on hegemonic stability theory, Keohane seeks to explain why such cooperation persists even after the decline of the hegemon's power relative to other states. While admitting that hegemonic leadership can be helpful in creating a stable

order wherein cooperation flourishes, he develops his functional theory of regimes to explain why such cooperation persists “after hegemony.” According to Keohane, even rational, egoistic states will have an incentive to participate in regimes because they help states overcome obstacles to achieving optimal outcomes. In this sense, international institutions promote cooperation between states because they help alleviate the problems associated with international anarchy: distrust and uncertainty between states and the transaction costs associated with interstate cooperation. In short, regimes are developed because actors in world politics believe that they help them make mutually beneficial agreements that would otherwise be difficult or impossible to attain (Keohane, 1984).

The Liberal Revival After the Cold War

The disintegration of the Soviet Union and the end of the cold war brought dramatic changes to domestic and international political structures and provided both challenges and opportunities for the various strands of both realism and liberalism. In a confident reassertion of the optimism and progressive outlook of liberalism, Francis Fukuyama’s (1989) essay, “The End of History?” proclaimed the ideological victory of liberalism over all other alternative theories of politics. For Fukuyama, the end of the cold war and the various democratic transitions in Africa, east Asia, Latin America, and eastern Europe represented the triumph of liberal capitalism and that there could be no improvement on its underlying principles and institutions, at least in theory. Fukuyama helped to revive the long-held view that the spread of liberal-democratic principles is the best prospect for a peaceful world order, spurring a vigorous scholarly debate on exactly how democracy, market economics, or both lead to peace, as well as the extent to which democracy needs to be consolidated within states in order for them to behave more peacefully.

Another related development in liberal thought in the post-cold war era was the proliferation of human rights norms, treaties and agencies, as well as a vigorous debate over what is known as *humanitarian intervention*. Although the principal human rights treaties predate the end of the cold war, the victory of the pro-Western forces in this ideological struggle made room for both state and nonstate actors to work more toward realizing human rights throughout the world. Although it is no coincidence that dominant international norms to a large extent reflect the values of the most powerful members of the international community, even though human rights are an essentially Western liberal idea, this idea has proven to be broadly appealing throughout the world—even in non-Western societies such as Japan and South Korea. Thus, the debate over human rights in international politics is not whether they exist or should be acknowledged, but rather when and how to implement them and how to

enforce these protections when states violate human rights (Forsythe, 2006).

A final development in liberal thought that gained prominence in the post-cold war era has been the rapid globalization of the world economy. Economic neoliberalism—a term generally used to refer to global market capitalism and free trade policies—has always favored the free play of market forces and the minimal role of the state in economic life. Yet liberal IR scholars view these developments in the context of the state and the international states system and focus on developments such as the growth of free trade, the increased ability of multinational corporations to escape states’ legal jurisdiction, the supposedly increasing irrelevance of state boundaries to the conduct of economic activity, and how these developments affect states’ behavior internationally (Friedman, 2000; Held, 1999). The idea of free trade and the belief in its efficiency and pacifying effects have nevertheless been the governing ideologies of the various free-trade institutions such as the WTO, EC, NAFTA, the IMF, and World Bank that have proliferated in the past two decades.

The Main Strands of Contemporary Liberal Theory

Liberal IR theory is a long and varied theoretical tradition that draws on some common foundational principles, insights, and ideas that in some way, shape, or form originated with European Enlightenment thought. But how can one make sense of or try to organize these different approaches? Although different books offer different categorizations, schools, or strands of liberal theory, there is some, but not universal, agreement on how to categorize the different approaches. This chapter offers four main categories, or strands, of contemporary liberal theory: pluralism, interdependence liberalism, institutional liberalism, and DPT.

The first strand of liberal thought is categorized here as pluralism, also known as *sociological liberalism* or sometimes *global governance theory*. Pluralism draws on the discipline of sociology to enhance state-centered approaches to IR by understanding relations between sub-state actors, or transnational relations—that is, individuals, groups, and organizations within states alongside traditional focuses on relations among political elites (Rosenau, 1980). According to Karl Deutsch (1957), increasing instances of transnational relations over time, and the increasing intensity of these interactions, can result in the creation of “security communities,” wherein potential and actual points of conflict between states can be addressed effectively, thereby promoting cooperation and peace. Thus, for pluralist theorists, transnational relations have the ability to not only facilitate cooperation by the presence of security communities, but can also foster the development of norms and rules promoting stability and peace in relations among states (Rosenau, 1990).

In short, for pluralists, IR is more than the study of relations among states and includes relations among private individuals, societies, and other groups. The more these non-state actors interact, network, and become interdependent, the less inclined their governments will be to resort to conflict.

Interdependence liberalism comprises the second grouping of liberal theory. Attempting to develop the earlier functionalist theory of David Mitrany (1966), Ernst Haas's (1958) groundbreaking neofunctionalist theory explained European integration in terms of political elites within Europe, identifying shared goals and interests and undertaking targeted cooperation on specific (economic) policy areas. Once integration of policy began, processes of cooperation and integration became self-reinforcing through the effects of spillover, whereby cooperation on a single policy necessarily leads to further cooperation on other policy areas in order to ensure policy effectiveness. Similarly, dealing with cooperation through institutions, Keohane and Nye's (1977) institutionalism bred the theory of complex interdependence, discussed previously. Complex interdependence details an international system where economic and social issues have become at least equal in importance to security concerns of and among states. Transnational relations therefore serve to transform a world based on political transactions occurring primarily between political elites to a world system where relations between influential citizens and nongovernmental organizations can wield significant influence on state actions as well. As such, international politics were transformed to appear to function in a manner similar to domestic political relations within states, thereby creating what has been termed *complex interdependence*.

More recent scholarship in this area can be found in the work of Anne-Marie Slaughter (2004) and focuses on what are known as *intergovernmental networks*. By looking within states at their different constituent institutions—that is, by disaggregating the state—Slaughter observes a complex web of networks between the various agencies of different states, such as law enforcement, environmental, financial, and a whole host of government agencies that are increasingly exchanging information and coordinating activity to address common problems on a global scale. Reminiscent of John Burton's (1972) "cobweb model" of transnational relations among private groups, Slaughter's networks are composed of government actors, which, unlike private actors, are more capable of being held accountable.

Institutional liberalism represents the third variant of liberalism and includes such approaches as regime theory and neoliberal institutionalism, which like interdependence liberalism, began with the observation that levels of international cooperation were much higher than could be explained by neorealists. The initial work that applied this institutional approach was that of regime theory, which focused not only on formal IOs but also on the broader

concept of *international regimes*, defined as sets of principles, norms, rules, and decision-making procedures in a given issue area (Krasner, 1983). This included not only IOs but also informal and nonbinding arrangements. Focusing on the prospects of both formal and informal international institutions for facilitating international cooperation, institutional liberalism argues that institutions are not merely weak tools of the state, but rather can provide vital channels through which cooperation can take place if states perceive the benefits of cooperation to outweigh the potential risks (Keohane, 1989). International institutions, organizations, and regimes help states overcome a range of collective-action barriers to cooperation by increasing opportunities and methods for information sharing, providing arenas for open discussion and negotiation between political elites and state actors and fostering a culture of cooperation. This not only makes cooperation easier by helping overcome collective action problems but also provides assurance and shared expectations that make defection from agreements more costly over time for all involved (Keohane, 1989; Keohane, Nye, & Hoffman, 1993).

The work of Robert Keohane discussed previously was seminal in the development of this approach in that it attempted to respond directly to the neorealists by accepting the neorealist assumptions that states are the dominant actors in international politics and that they are rational actors—that is, that they calculate the costs and benefits of certain actions and take the action that gives them the highest payoff. This led some to accuse the neoliberals of essentially being neorealists in disguise, except with a focus on international (economic) institutions. Yet there are some subtle, yet important, distinctions. According to David Baldwin (1993), neorealists and neoliberals first disagree on the nature and consequences of anarchy, with neorealists seeing anarchy as placing more severe constraints on states than neoliberals. As a corollary, neorealists view international cooperation as both more difficult to achieve and maintain and more dependent on state power, and neorealists are therefore more skeptical regarding the ability of institutions to mitigate anarchy. Second, neorealists assume that states are more concerned about relative gains, whereas neoliberals have emphasized absolute gains. In other words, when states are faced with the possibility of cooperating for mutual gain, neorealists argue that states will be concerned how much they will gain vis-à-vis other states, whereas neoliberals believe states are concerned primarily with their own gains and are largely indifferent to the gains of other states. Third, neorealists and neoliberals differ on the priority of state goals in that the former emphasize security and survival, and the latter, following the interdependence liberals, argue that states are more concerned with economic welfare. Finally, the two schools differ on threat perception. Whereas neorealists assume that a state's capabilities or power is decisive in how the state will behave, neoliberals argue that it

is not just the capabilities of a state that matter, but also the state's intentions. Thus, neoliberals emphasize intentions, interests, and information as explanatory variables, whereas neorealists emphasize the distribution of capabilities (Baldwin, 1993).

The last strand of liberal theory this chapter discusses is DPT, also commonly referred to as *republican liberalism*, which has been given considerable prominence among liberal theories in the contemporary study of IR. Essentially, democratic peace theorists, including the early writings of Kant, observed that democratic states do not go to war with one another. Reasons given for the peaceful relations existing among democratic states include the argument that since democratic governments are answerable to their citizens, the risk of electoral ramifications for leaders undertaking war with another democratic state is fairly high; the observation that democratic societies tend to value peaceful resolution of conflict; and finally, the empirical observation in the tradition of the previous three strands of liberal theory that democratic states tend to be highly interdependent on one another through membership in international organizations, institutions, and regimes (Doyle, 1997; Gilpin, 1981; Lipson, 2003; Russett & O'Neal, 2001). Furthermore, the dramatic increase in the number of democratic and democratizing states from the 1970s to the 1990s, in what has been termed the *third wave* of democratic development, provided increased salience to DPT and has contributed to the prominence of this approach in the study of IR (Huntington, 1991).

Conclusion: Applications and Challenges

With the end of the cold war, the continuing globalization of the world economy, and the atrocities caused by global terrorism, the traditional issues that occupied the liberal research agenda have been endowed with a new sense of urgency. Issues of trade and global economics remain central to research agendas of institutional and interdependence liberals, particularly the study of international organizations, both new and old. Since the collapse of the world economy in late 2008, there has been an urgent need for more knowledge on how IOs such as the WTO, the Bretton Woods institutions, and the various free-trade organizations such as the EU and NAFTA have either contributed, or may be used as a solution, to the current financial crisis. Likewise, research on international networks of banking and other international financial institutions requires further development as the processes of integration and transnational relations continue to intensify in places like western Europe.

Furthermore, especially since the 2003 Iraq War, there has been a renewed debate over the democratic peace and calls for more research on exactly how democracy leads to peace and whether or under what conditions it may be permissible to forcibly change the government of a state to

make it more democratic and peaceful. The role of IOs in the area of international security is also a pressing concern, as NATO evolves from its cold war posture into a tool for democratic enlargement and an entity better equipped to deal with terrorist and insurgent challenges in places like Afghanistan. Newer institutions such as the African Union are likewise increasingly becoming the subject of analyses regarding how this entity, in conjunction with the United Nations, can become more effective at addressing the numerous crises on the African continent, such as in Darfur, Sudan, and the Democratic Republic of Congo. Likewise, the potential emergence of nuclear threats from Iran and North Korea has made collective efforts at nuclear nonproliferation a particularly important subject as states attempt to use the international nonproliferation regime to prevent the spread of weapons of mass destruction. Such questions that deal with the emergence, change, and effects of IOs therefore remain crucial to the institutional and interdependence liberal research agendas.

Finally, the attention given by governments to nonstate terrorism and the resources dedicated to combating it demonstrate an urgent need for more research by pluralist liberals regarding the threats of nonstate armed groups, including global terror networks like al Qaeda, as well as more regional groups such as Hezbollah. What is the best way for states to cooperate to combat this threat? What is the relationship between liberal democracy and terrorism? The current wave of Islamic militancy is profoundly antiliberal and therefore presents a threat not only to liberal states but also arguably to the global order over which liberal states have presided. Although some have argued that the emergence of al Qaeda and its affiliates is evidence of the deterritorialization of international politics and the further decline of the sovereign state, others argue that it has allowed the state to accumulate more power, including placing new restrictions on civil liberties, enhancing powers of surveillance and detention, and increasing military spending (Harvey, 2003). As Scott Burchill (2005) notes, the threat posed by Islamic terror has been met by an increase in military activity by powerful states that have been emboldened to intervene—even preventively—in other states' internal affairs.

Liberalism as portrayed in this chapter is an inherently optimistic approach to understanding international relations that emphasizes the role of international institutions, free trade, domestic (liberal) political institutions, and non-state actors as all having important influence on international politics. Virtually all liberal scholarship is imbued with a faith that there can be progress in human affairs. Although liberalism may have originated as a broad philosophical statement about human progress and perfectibility, it is today best understood as an analytical project concerned with exploring the possibilities for international peace and cooperation and for improving the human condition (Sterling-Folker, 2006). Although there has been reason to be optimistic about the outlook for international

political life over the past decades—particularly since the end of the cold war—recent years have witnessed profound changes that continue to challenge this optimistic outlook. The task for liberal IR scholars today is to improve our knowledge of these various changes in order to gain a better understanding of their causes and consequences in the hopes that they can be better understood and ultimately overcome.

References and Further Readings

- Angell, N. (1913). *The great illusion: A study of military power to national advantage*. New York: Putnam.
- Baldwin, D. A. (1993). *Neorealism and neoliberalism: The contemporary debate*. New York: Columbia University Press.
- Boyle, M. (2004). Utopianism and the Bush foreign policy. *Cambridge Review of International Affairs*, 17(1), 81–103.
- Burchill, S. (2005). Liberalism. In S. Burchill et al. (Eds.), *Theories of international relations* (pp. 55–81). New York: Macmillan.
- Burton, J. (1972). *World society*. Cambridge, UK: Cambridge University Press.
- Carr, E. H. (2001). *Twenty years' crisis, 1919–1939: An introduction to the study of international relations*. New York: Palgrave.
- Deutsch, K. W. (1957). *Political community and the north Atlantic area*. Princeton, NJ: Princeton University Press.
- Doyle, M. W. (1997). *Ways of war and peace: Realism, liberalism, and socialism*. New York: W. W. Norton.
- Forsythe, D. P. (2006). *Human rights in international relations* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Friedman, T. (2000). *The Lexus and the olive tree*. New York: Farrar, Straus & Giroux.
- Fukuyama, F. (1989). The end of history? *The National Interest*, 16, 3–18.
- Gilpin, R. (1981). *War and change in world politics*. New York: Cambridge University Press.
- Haas, E. B. (1958). *The uniting of Europe: Political, social and economic forces 1950–1957*. Palo Alto, CA: Stanford University Press.
- Harvey, D. (2003). *The new imperialism*. Oxford, UK: Oxford University Press.
- Held, D. (Ed.). (1999). *Global transformations*. Cambridge, UK: Cambridge University Press.
- Howard, M. (1978). *War and the liberal conscience*. New Brunswick, NJ: Rutgers University Press.
- Huntington, S. (1991). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Hutchings, K. (1999). *International political theory*. London: Sage.
- Jackson, R., & Sorensen, G. (2007). *Introduction to international relations: Theories and approaches*. Oxford, UK: Oxford University Press.
- Kant, I. (1991). Perpetual peace: A philosophical sketch. In H. S. Reiss (Ed.), *Kant's political writings* (2nd ed., pp. 93–130). Cambridge, UK: Cambridge University Press.
- Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton, NJ: Princeton University Press.
- Keohane, R. O. (1989). *International institutions and state power: Essays in international relations theory*. Boulder, CO: Westview Press.
- Keohane, R. O., & Nye, J. S. (Eds.). (1971). *Transnational relations in world politics*. Cambridge, MA: Harvard University Press.
- Keohane, R. O., & Nye, J. S. (1977). *Power and interdependence: World politics in transition*. Boston: Little, Brown.
- Keohane, R. O., Nye, J. S., & Hoffman, S. (Eds.). (1993). *After the cold war: International institutions and state strategies in Europe, 1989–1991*. Cambridge, MA: Harvard University Press.
- Knutsen, T. L. (1997). *A history of international relations theory*. New York: Manchester University Press.
- Krasner, S. D. (1983). *International regimes*. Ithaca, NY: Cornell University Press.
- Lipson, C. (2003). *Reliable partners: How democracies have made a separate peace*. Princeton, NJ: Princeton University Press.
- Locke, J. (1980). *Second treatise of government* (C. Macpherson, Ed.). Indianapolis, IN: Hackett.
- Long, D., & Wilson, P. (Eds.). (1995). *Thinkers of the twenty years' crisis: Interwar idealism reassessed*. Oxford, UK: Clarendon.
- Miller, J. D. B. (1995). Norman Angell and rationality in international relations. In D. Long & P. Wilson (Eds.), *Thinkers of the twenty years' crisis: Interwar idealism reassessed* (pp. 100–121). Oxford, UK: Clarendon.
- Mitrany, D. (1966). *A working peace system*. Chicago: Quadrangle.
- Morgenthau, H. (1993). *Politics among nations: The struggle for power and peace*. New York: McGraw Hill.
- Ricardo, D. (1911). *The principles of political economy and taxation*. London: Dent.
- Rosenau, J. N. (1980). *The study of global interdependence: Essays on the transnationalization of world affairs*. New York: Nichols.
- Rosenau, J. N. (1990). *Turbulence in world politics: A theory of change and continuity*. Princeton, NJ: Princeton University Press.
- Russett, B., & O'Neal, J. R. (2001). *Triangulating peace: Democracy, interdependence, and international organizations*. New York: W. W. Norton.
- Slaughter, A. M. (2004). *A new world order*. Princeton, NJ: Princeton University Press.
- Smith, A. (1993). *Wealth of nations* (K. Sutherland, Ed.). Oxford, UK: Oxford University Press.
- Sterling Folker, J. (Ed.). (2006). *Making sense of international relations theory*. London: Lynne Rienner.
- Vattel, E. (1863). *The law of nations* (J. Chitty, Trans.). Philadelphia: Johnson.
- Waltz, K. (1979). *Theory of international politics*. Reading, MA: Addison Wesley.
- Wilson, P. (1995). *The twenty years' crisis and the category of "idealism" in international relations*. In D. Long & P. Wilson (Eds.), *Thinkers of the twenty years' crisis: Interwar idealism reassessed* (pp. 1–24). Oxford, UK: Clarendon.

39

DEPENDENCY AND WORLD-SYSTEMS

GLORIA C. WALKER

Centenary College of Louisiana

The economic and political instability faced by former colonies and other developing countries in the post–World War II era sparked a great amount of concern and much debate within policy circles and scholarly research. Within political science, the resultant research became known as *political development* and encompassed the fields of comparative politics, international relations, and international political economy. One of the main concerns of this literature is the unequal and inequitable economic and political development that exists between developing and developed countries. In response, two main schools of thought emerged, modernization and dependency (of which world-systems is a part), each taking a very different approach to explaining the origins and effects of this observed lack of development. Before describing dependency theory, however, it is first necessary to put this work in context by describing modernization theory, which emerged first and to which dependency was a direct response. Since modernization theory is discussed in detail elsewhere in this handbook, only a brief introduction is provided below.

The primary assumption of modernization theory is that there is one path to development, which consists of adopting the modern values, behaviors, technology, and institutions of developed countries. The focus of the modernization approach is thus to apply the so-called Western (western Europe and the United States) development experience to the rest of the world. For development to occur, a country

must replace traditional norms and structures such as rural society, rigid social structures, passive citizen participation, undifferentiated political structures, traditional and elitist sources of authority, subsistence agriculture, and primary economic activity with modern structures and values including geographic mobility, urbanization, flexible social structures, active citizen participation, complex political structures, merit-based authority structures, formalized judicial system, rule of law, and economic diversification. The success with which this diffusion of modern values and institutions occurs determines the eventual political and economic development achieved. Once traditional values and institutions are replaced with modern versions, there is no turning back since such changes cannot be undone.

Modernization theory thus describes a linear and irreversible process that creates conditions conducive to economic growth and political development. It is also an internally driven process that gives no consideration to external or international influences that may adversely affect a country's ability to develop effective political and economic structures. If a country fails to adopt modern ways, modernization theory concludes this is due to inferior or so-called wrong values. Since culture and values are viewed as deeply rooted and difficult to change, a country with the wrong values is unlikely to benefit from any guidance and assistance provided by developed countries, and thus their policymakers may conclude that aid and investment will not contribute

to modernization and redirect their money to countries with the so-called right values.

The assumption that both a single linear path to development and the failure to develop are due solely to internal factors was viewed by many scholars, especially those from developing countries, as being an incomplete, value laden, and offensive approach to development. In response, scholars, economists, and policymakers from Latin America, known as *dependentistas*, took exception to modernization theory's Eurocentric approach and internal focus and developed dependency theory, an externally oriented argument, to explain the lack of political and economic development in developing countries.

Theory

Dependency and world-systems theory are often presented as unified approaches to development since both are primarily concerned with the ways in which external factors, particularly the international capitalist economic system, have affected and continue to affect developing countries. They did, however, emerge separately, with the former originally focusing on the political economy of Latin American underdevelopment and the latter building on dependency by placing it within a larger global historical context. Dependency-based work began to appear shortly after World War II but reached its peak in the late 1960s and 1970s when modernization had established itself as the dominant approach to development and as increasing concern arose in academic and policy circles regarding the fate of former colonies. World-systems emerged in the 1970s as a historical-sociological approach seeking to place dependency in historical perspective by focusing on what are termed *long cycles* in history and the effects of the most recent cycle on developing countries.

Both approaches are also often identified with Marxism since they argue that imperialism, exploitation, and the international capitalist economic system are to blame for underdevelopment. Imperialism and exploitation are key concepts for dependency and world-systems, but instead of applying them internally to a domestic class-based system as with Marxism, they are applied externally to the international economic system, which emulates a class system at a global level. Another parallel is the argument that to escape exploitation, revolution must occur. In Marxism, the proletariat must overthrow the bourgeoisie, whereas in dependency the countries of the periphery (developing countries) must revolt against economic domination by the core (developed countries); however, the timing of revolution is a key point of disagreement. Marx and Engels (2002) argued in *The Communist Manifesto* in 1848 that all countries go through linear stages of development, with socialism being the last and highest level thereof. Most dependentistas, however, find Marxism's linearity

unacceptable and argue that rather than incurring long-term suffering while waiting for the right point in the process at which to have their revolutions, developing countries need to bypass the stages and take immediate action to improve their situations.

The first writings in what would become dependency theory were from neo-Marxist Latin American economists and sociologists seeking to explain why, despite more than 100 years of independence, Latin America continued to lag far behind western Europe and the United States in terms of economic development. An early and influential work was produced in 1950 by Raul Prébisch while he was working for the United Nations Economic Commission for Latin America (ECLA). It was in this groundbreaking work that the concepts of core (or center) and periphery were introduced along with the assertion that the core exploits the periphery. The arguments developed in this work sparked what would become a vibrant and lively debate between modernization theorists and dependentistas.

Dependency

In response to modernization's claim that to stimulate economic growth, developing countries simply need to duplicate the path of the developed countries, dependency argues this is impossible since that path no longer exists. The international capitalist economic system did not exist when the West began the development process; instead, the system is a product of the West's development that was created to facilitate its continued economic growth. Developed countries were in fact able only to grow their economies, industrialize, and accumulate significant wealth through the political and economic exploitation of other countries by way of colonialism and imperialism. Although colonialism has ended, it remains in developed countries' best interests to maintain this system given the advantages it provides, and as a result, they actively prevent developing countries from growing competitive economic systems to ensure their continued exploitation. As a result, the rules of the development game are very different for developing countries since they are faced with a coherent and exploitative international economy that is structured to keep them at a disadvantage.

Dependency theory thus argues that rather than being cultural, development is in fact structural and dependent on the nature of a country's insertion into the international capitalist economic system. For developed countries, this insertion was voluntary and purposeful since they created the system for their own benefit. For former colonies, however, insertion was originally involuntary via colonialism, and on gaining independence, the new states were weak and uncompetitive. The prospects for development are thus influenced by the long-term political and economic effects of the state's insertion into this system.

Early Dependency Theory

Dependence is most clearly defined in Theotonio Dos Santos's (1970) article "The Structure of Dependence":

By dependence we mean a situation in which the economy of certain countries is conditioned by the development and expansion of another economy to which the former is subjected. The relation of interdependence between two or more economies, and between these and world trade, assumes the form of dependence when some countries (the dominant ones) can expand and can be self sustaining, while other countries (the dependent ones) can do this only as a reflection of that expansion, which can have either a positive or a negative effect on their immediate development. (p. 231)

Dependence thus involves an asymmetrical power relationship between the core and the periphery, in which the latter has little or no ability to grow and expand because of its continued exploitation by the core. This asymmetrical relationship can ultimately be traced to colonialism since it is the structures put in place during this era that continue to limit development opportunities in the periphery today.

Under traditional colonialism, conquest and subjugation created conditions for the monopolization and exploitation of a colony's natural resources (including labor). This, in turn, facilitated the rapid development of core economies by allowing them to cheaply fuel industrialization, develop economic diversification, and effectively compete and trade with each other. The effect on the colony was to make it a provider of cheap raw materials and agricultural products to be exported to the colonial power. As a result, the colony's survival was dependent, in the truest sense of the word, on the colonial power. Although colonialism via political domination officially ended with decolonization, dependency asserts that the core still economically dominates and exploits the periphery, a condition referred to by dependentistas as *neocolonialism*. Despite the name, neocolonialism applies to all countries in the periphery, not just former colonies, since the structure of the international capitalist economic system creates the same disadvantages for all developing countries.

Under neocolonialism, the economic situation for the periphery remains relatively unchanged as the ability of the newly independent countries to expand and diversify their economies is limited by the core's continued control of the global economic structures put in place during the colonial era. The periphery remains limited primarily to the production and export of agricultural goods and raw materials and thus finds itself having to import finished goods, technology, and other industrial goods from the core. The wealth generated in the international system continues to flow primarily to the core since, as the financier of the production process, it receives the profit from the finished product as well as any taxes from its sale. The periphery, on the other hand, receives an artificially low

agreed-on price for its raw materials and agricultural products (because of the core's superior negotiating power) with very little, if any, money reinvested in its economies and production processes. Given the relatively high value of the goods the periphery imports from the core, the relatively low value of the goods it exports to the core, its minimal industrial capacity, and its low ability to diversify, the periphery receives a disproportionately small share of gains from trade, has generally unfavorable terms of trade, and builds increasing trade deficits.

Another issue related to colonial legacy and neocolonialism is the need for countries in the periphery to secure foreign investment and borrow capital from the core in an attempt to overcome their structural disadvantage and compete effectively in the international economy. Foreign investment brings desperately needed capital and industry to the periphery but ultimately extracts more than it adds as profits are returned to and reinvested in the core, not in the periphery. The ability to borrow money as a means of raising investment capital is generally considered to be beneficial to developing countries, but dependentistas argue that international financial institutions such as the International Monetary Fund and the World Bank create a perpetual cycle of debt through high interest rates and predatory lending practices. The core, for all practical purposes, thus directs the most profitable sectors of the dependent country's economy as well as state economic policy and is, as a result, able to keep the periphery dependent on it for infusions of capital. The negative effects of foreign direct investment and debt are considerable for developing countries and leave them at a severe competitive disadvantage and in a perpetual state of underdevelopment. Rather than being a means for escaping poverty and improving economic performance, foreign capital in reality structurally limits growth potential, with the capitalist system serving as another imperialist tool to maintain control of the periphery and to ensure that the periphery remains dependent on the core.

Although the process just described is economic in nature, the economic control the core wields gives it strong de facto political control over the periphery. This continued political dependence arises from the incentive structure created under neocolonialism and the need for the periphery's political, economic, and military elite to maintain dependent economic relationships for fear the periphery will lose power if it defies the existing structures and agreements. Members of the periphery elite thus find themselves supporting a system that closely resembles the exploitative colonial structure in which only a powerful few benefitted, except now the benefits also accrue to them. The incentive structure is ultimately not conducive to providing collective gains for the entire population but rather selective gains for the elite. As a result, although the core may no longer have the direct political control present under colonialism, it still has significant indirect control leading to continued political dependence.

The prospects for economic growth and political development in dependent countries are thus grim. Since developing countries are relegated to an exploitative system in which production is limited to raw materials and unfinished goods, ruled by elites with a vested interest in maintaining the status quo, and dominated by a core with a vested interest in ensuring it stays that way, development is unlikely. Thus, dependentistas argue, if countries in the periphery cannot develop while participating in the international capitalist economic system, they must withdraw from it. In essence, if the rules of the game are biased against you, then you have two choices: to continue to play by the rules and thus continue to be exploited or to rewrite the rules in a way that does not leave you at an unfair disadvantage. Absent significant changes, the periphery is essentially doomed to continued economic and political dependence. The solution is to remove the periphery from the exploitative capitalist system by creating a competing economic order either with or without full-scale socialist revolution.

Dependent Development

In the late 1960s and early 1970s, some Latin American countries (Argentina, Brazil, Colombia, and Mexico in particular) began to experience economic development through diversification and industrialization that was largely fueled by increased investment from multinational corporations (MNCs). This naturally called into question some of the assumptions of dependency theory—particularly the role of foreign investment—and its pessimistic conclusions regarding the prospects for development. Fernando Henrique Cardoso and Enzo Faletto argued in 1969 (the publication date in the original Spanish) that the nature of dependency had changed for these countries because they possessed more resources and had higher industrial capacity than other dependent countries and as a result were able to gain more leverage against and reduce the constraints imposed on them by the core. They were no longer fully dependent, but were instead in a state of “dependent development,” an economic status in which there is growth and some measure of increased economic control, but ultimately, the core maintains significant influence over policy and development (Cardoso & Faletto, 1979).

Dependent development is the result of the internationalization of a dependent country’s internal market via MNCs. Given the profit imperative of corporations, many have sought to cut costs and increase profits by shifting some of their production capacity to developing countries where the cost of land, labor, and raw materials is lower than in the corporation’s home state. The MNC thus negotiates with the political, economic, and perhaps even military elite in the host country (the dependent state) to allow it to open a factory, mine for minerals or metals, or refine a raw material. In exchange, the host country benefits from taxation and fees (usually negotiated at a reduced rate) from

the MNC’s activity, diversification of the economy, lower unemployment, and in some cases, improved infrastructure in the form of roads, ports, railroads, or airports that results from the MNC’s need for better access to transportation.

On the surface, MNC investment appears to be beneficial to the host country, but there are other factors to consider that, according to dependent development, may lead to economic decline, political instability, or both. MNCs invest in periphery countries only when it helps to maximize their profits, and thus although the internationalization of a host country’s economy results in increased development, this development is ultimately only a side effect of an MNC’s investment. All profits, aside from those needed to maintain the offshore facility, are invariably invested back into the MNC, thus profiting the home country through tax revenues it receives from the MNC’s domestic and international activities. In addition, all decisions regarding MNC activity are made at corporate headquarters, and thus, key economic decisions that could adversely affect the host country are made without its input and out of its control. Should the MNC’s profits decline or disappear, it has no incentive to remain in the host country and is thus likely to withdraw and seek a better deal in another developing country. The results of MNC withdrawal are potentially devastating since without the MNC driving development, the economic situation is likely to deteriorate, leaving many local workers unemployed and eliminating the economic benefits the host country was receiving.

Dependent development also has implications for political stability within the host country because of the uneven distribution of wealth and increased income inequality that arises from economic growth. The elite benefit greatly from their arrangements with the core, while the masses receive little to no relief from extreme poverty. This in turn has the potential to lead to civil unrest and domestic conflict as the masses challenge the status quo, demanding improved living conditions and reduced income inequality. These circumstances put the elite in an untenable position as they must choose between continued development and domestic turmoil, with the latter potentially leading to open rebellion or even revolution. The most likely outcome in this situation is for the state to implement coercive measures against the masses to ensure the development agenda is not compromised. The MNC–elite alliance thus often serves to develop and maintain repressive governments (Evans, 1979).

Although under the conditions of dependent development countries are able to develop beyond those experiencing full dependence, the process is still neocolonial in nature. Profitable economic sectors remain under the control of the core, except rather than the core state having direct control, MNCs, acting as their agents, now fulfill this role. The host state, on the other hand, experiences growth through foreign investment, but this growth is ultimately dependent on maintenance of the exploitative

relationship, leaving the developing country vulnerable to changing economic conditions in the core. In the end, the relationship remains heavily skewed in favor of the core and structured in a way that keeps the periphery at a competitive disadvantage.

World-Systems

World-systems theory was inspired by dependency's arguments and is most closely identified with Immanuel Wallerstein (1974). Although world-systems has its origins in historical inquiry, theorists argue that the approach in fact transcends traditional disciplinary boundaries since it considers the world as a whole, not just from a political, economic, sociological, or historical standpoint. As a result, world-systems theorists distinguish between systems of the world, as represented by the focus on discrete units such as states, and systems that are a world, represented by a system that is larger than any single national or political unit. A world-system does not necessarily encompass the entire world, but it does integrate multiple political and cultural units whose behavior and interactions are guided by a set of systemic rules and cultural norms known as a *geoculture*. There are two types of world-systems: world-empires and world-economies. In world-empires, there is a single political authority for the entire system and a common culture that often dominates peripheral areas. In world-economies, on the other hand, although there is no single political authority, there is a dominant economic structure that guides state interactions and creates an integrated system of production that results in a division of labor based on industrial capacity.

The modern world-system, which is a world-economy, originated in the so-called long 16th century in western Europe, which began with the Spanish discovery of the Americas in 1492 and ended with the Peace of Westphalia in 1648. During this period, the core of economic power shifted from the old Mediterranean to northwest Europe. As part of this shift, the old core (the Mediterranean) became the semi-periphery (a new category of developing country described subsequently). The rest of the world that was economically relevant at the time, Latin America and eastern Europe, then became the periphery. The French Revolution of 1789 was another crucial point for the current world-system since it solidified liberalism as the dominant philosophy of the core and thus implemented a new geoculture. With liberalism came the concept of universalism, an important principle stressing equality in citizen access to the system and in the application of laws. Wallerstein (1974) notes that although these values are generally common to core countries, they are not necessarily applied and encouraged elsewhere, leading to inequality within the system. The completion of the modern world-system occurred in the late 19th century when the rest of the world, previously external to the system, became economically integrated into the world-economy via colonization and industrialization's need for more and

cheaper resources. In focusing on long cycles of economic and political development, world-systems broke with the dominant perspective that the modern world order was a product of World War II and its aftermath and argued instead that it was in fact 500 years in the making.

The modern world-system added a new category of state, the semi-periphery, to the relationship between the core and the rest of the world. The semi-periphery consists of countries that possess both core- and periphery-style production capacities. The core-type industries produce higher cost, higher value manufactured goods and behave no differently than producers in core countries in that they exploit countries in the periphery. The periphery-type industries, on the other hand, produce lower cost, lower value goods, such as raw materials and agricultural products, and are exploited by core countries. Countries in the semi-periphery thus have a mix of production capacities that creates two very different economic sectors within a single country and as a result makes it vulnerable to domestic conflict: Those tied to core-type industry seek to maintain what they have at all costs, and those tied to periphery-type industry seek to improve their situations. This is essentially the same relationship one sees between core and periphery countries, except at work within a single country. The relative mix of core- and periphery-type industries determines the degree to which a state is ultimately still dependent on the core and whether it escapes the periphery to join the semi-periphery. However, despite the fact that countries in the semi-periphery are in a better situation than those in the periphery, they remain far less developed than countries in the core and because of the structural constraints of the world-economy will not be able to escape semi-periphery status.

The way in which the state and political power are treated analytically differs between world-systems and dependency. Dependency is a state-oriented approach in terms of the primary level of analysis, yet it tends to ignore state power as an important factor in explaining variation in development levels across countries and relies instead almost exclusively on the international capitalist economy to explain the differences. The world-systems approach, on the other hand, argues that the appropriate level of analysis is the world-system since the study of discrete units such as states, their national histories, and national political structures is analytically limiting given the global nature of state activity. However, world-systems theory recognizes that states are capable of exercising considerable political power and thus contribute to explaining variation in development levels. In fact, state power plays a key role in world-systems since international political history is replete with examples of one state after another upsetting the system in an attempt to gain more power and, ideally, control of the world-system itself. State power is thus considered a key component in explaining the historical emergence of the core and its ability to subjugate and exploit the periphery. State power, or rather lack thereof, also helps to explain the inability of the periphery to

develop since these states are either in a colonial situation in which they completely lack autonomy or in a neocolonial situation in which autonomy is low and external influence and control are high.

The effect of this world-system has thus been to create a world-economy with a reach never before seen. It is based on capitalist economic principles with its primary priority being the accumulation of capital for its own sake. To achieve this priority, the system is characterized by a worldwide division of labor in which economic structure (core, periphery, semi-periphery) is determined by a country's or region's specific mix of economic activity and each zone being rewarded differently by the world economy, with the core rewarded with the highest levels of surplus and income. The modern world-system is thus ultimately an unequal system in which the core, because of its superior wealth and state power, is able to exploit the periphery and semi-periphery to ensure its own continued enrichment and development.

Wallerstein (2004) later updated his argument to include what he refers to as the current crisis of the modern world-system. This crisis began with the cultural shocks of 1968, which were characterized by widespread social upheaval in the core, periphery, and semi-periphery as a result of increasing disillusionment with the promises of capitalism and democracy's apparent inability to adequately and appropriately foster economic development, political stability, and social equality. In terms of world-systems theory, many in the periphery and semi-periphery no longer believed investment by the core was the answer to development but instead simply a means of maintaining the core's power, and thus, large numbers of the exploited and their sympathizers no longer believed that if they were patient, the system would reward them (as it did the core).

These changing beliefs about the nature of the world-system have thus led to increasing questioning of and agitation against the existing system and form the basis of a large-scale rejection of the dominant liberal geoculture that has been emergent since the French Revolution. The result, which jeopardizes the capitalist world-economy, is a destabilization of the world-system, which Wallerstein (2004) argues is reflected in increased institutional instability and political violence throughout the periphery as well as reduced cooperation with international financial institutions such as the World Bank and the International Monetary Fund. The effects of 1968 are thus still felt today and, according to Wallerstein, are likely to continue to be felt for another 25 to 50 years while the existing world-system either stabilizes or a new system emerges.

Critical Evaluation and Empirical Evidence

Early work in dependency was primarily theoretical and anecdotal, and given the highly controversial nature of the theory, it has been subjected to strenuous critical evaluation

and empirical testing. Critical evaluation has occurred on both theoretical and methodological grounds and points out a number of key weaknesses of dependency. Empirical analysis has been undertaken by both proponents and critics seeking to empirically test dependency's assumptions and implications and has yielded mixed results. Ultimately, although there is some support for dependency, there is not clear and consistent evidence that satisfies its critics.

Critical Evaluation

Dependency theory has a number of recurring criticisms. First, the theory is accused of being both too broad and too narrow. It is too broad because it seeks to be a universal explanation in attempting to explain all cases of a phenomenon while using a single set of assumptions and variables. In the case of developing countries, underdevelopment can be explained for all of them with the same cause and process—namely, exploitation via position in the global economic order. Dependency theory is also accused of having too narrow a focus given the argument that the capitalist international economic system (a single external factor) is the cause of dependence and underdevelopment without taking into consideration important internal factors such as the role of class, culture, state strength, or ethnic heterogeneity. Some dependentistas went so far, for example, as denying tribal relations and conflicts in precolonial Africa a role in the postcolonial political and economic situation even after old ethnic tensions erupted into violence. In treating local history and actors as irrelevant to a country's development, dependency theory ignores potentially contradictory evidence, thus leading to research results that may be skewed toward supporting the theory despite the existence of information that could in fact dispute it (Smith, 1979).

Another critique is that the concept itself is vague, ill-defined, and potentially tautological because the definition generally includes the very concepts that need to be defined, specifically, exploitation and underdevelopment. Furthermore, dependence is not appropriately operationalized, meaning the definition leads to measuring it in a way that limits its analytical power. Dependence is typically treated as a dichotomy with a country categorized either as dependent or not dependent, rather than as a continuum in which there are various degrees of dependency. Just as there are great disparities between developed countries, there are also great disparities between dependent countries. As a result, the effects of dependency are likely to vary depending on the relative dependence of each individual state. Given this, critics ask whether the same set of policies or solutions is appropriate for all developing countries or whether there is a need to differentiate policy by degree of dependence. In treating all dependent countries the same way, dependency theorists potentially compromise any inferences and conclusions they may draw from their analyses.

The final main critique is that given the primarily anecdotal and historical evidence used by dependentistas, especially in early writings, the evidence is often thin in that it focuses on establishing historical processes without applying empirical tests of actual conditions within countries that would support the theory. Baran (1957), for example, argues that were it not for British exploitation, India would have had a much smoother and much less traumatic development process, yet how can one know this? How does one test for what might have been? Arguments based on conjecture are ultimately untestable since one cannot gather nonexistent empirical evidence. This was in fact a common pitfall of much early dependency literature and became a major issue resulting in a flurry of empirical tests, a summary of which is provided in the next section.

Empirical Evidence

Empirical analyses of dependency theory have been conducted by both proponents and critics, with the former attempting to develop solid evidence supporting dependency's assumptions and implications and the latter seeking to determine whether there is reliable support for the theory's universalist arguments. Tests of dependency apply a variety of methodological approaches including single case studies, regional studies, cross-regional comparisons that incorporate cases from two or more regions, historical analysis, and statistical analysis. A key issue with empirical analyses of dependency is the difficulty in forming testable hypotheses given the historical and anecdotal approach discussed earlier. Another problem is that, depending on the part of the theory being tested or the country or region being examined, different variables and measures are used, which can make it difficult to compare results across studies and develop a coherent and consistent body of evidence. Ultimately, despite the various approaches used to test dependency, empirical testing of the theory has yielded mixed results, with some studies claiming a relationship between dependency and underdevelopment (Bradshaw, 1985; Chase-Dunn, 1975) and others failing to find sufficient evidence to support the theory's contentions (Jackman, 1982; McGowan & Smith, 1978).

In empirically comparing developing countries, a number of exceptions, particularly the rapid and sustained economic growth and relatively equitable income distributions of the so-called Asian Tigers (South Korea, Taiwan, Hong Kong, and Singapore) from the 1960s to the mid-1990s, call into question dependency theory's claims regarding the inability of periphery or semi-periphery countries to experience significant economic growth and development. These countries in fact achieved sustained growth through close ties with the core, consisting of a combination of trade, loans, foreign investment, and technology transfers, all of which are factors dependency theory argues will inhibit growth and deepen dependence. Despite the 1997

economic crisis that appeared temporarily to validate this argument (since the area was heavily dependent on foreign investment), the region has experienced a solid recovery, which would again tend to dispute the conclusions and implications of the dependency approach. The shifting of European Mediterranean countries from core to semi-periphery and back to core again also appears to defy the predictions of dependency theory. All of these countries have apparently found a way to escape dependent status, something dependency theory argues is extremely unlikely. Critics of this approach point to such cases as evidence of its limited explanatory power and argue the reason for this is the failure of dependency theory to take internal factors into consideration.

Overall, the evidence tends to show that those developing countries experiencing the highest economic growth rates, rising per-capita income, and reduced income inequality (the Asian Tigers, for example) have very close ties to the international economic system and to the core. In general, countries with higher levels of foreign investment and lower trade barriers (tariffs, regulations, etc.) experience high levels of growth at a rapid rate. Those applying significant trade barriers and with lower levels of foreign investment, on the other hand, tend to be furthest removed from the core and among the poorest of the developing countries. These relationships run counter to dependency's predictions. However, as dependency theory predicts, countries that rely primarily on the export of raw materials and agriculture do tend to be very poor and not perform well in terms of economic and social development indicators.

Policy Implications

Dependency and world-systems theory sought to shift the focus of policymakers away from the assumptions and implications of modernization theory and toward those of dependency, particularly the issues of exploitation by the core, predatory lending practices by international financial institutions, and the negative effects of relying on foreign aid and investment.

A popular approach for developing countries seeking to improve their conditions and escape dependence was to implement import substitution industrialization (ISI) policies. These policies have three main characteristics: promotion of domestic production of industrial goods that substitute for imported goods, an overvalued currency that allows the emerging manufacturers to import the heavy equipment and other machinery needed for production, and protectionist barriers to trade such as tariffs on imports intended to protect the developing country's infant industries from foreign import competition. As a result, these policies are intended to replace foreign control of the economy with domestic production and thus result in the discouragement of foreign direct investment, particularly by

MNCs. The logic is that in substituting domestically produced versions of products for imports from the core, the country becomes more self-sufficient and thus reduces dependence. As a result, the policy partially implements dependency's recommendation that dependent countries must distance themselves, if not remove themselves outright, from the international economic system.

One result of this policy was increasing budget and trade deficits. Budget deficits emerged as government spending increased to develop the infant industries and build infrastructure to support them without having the necessary tax revenue to pay for it. Trade deficits emerged because of the overvaluation of currency, which resulted in lower exports as the developing country's products became increasingly uncompetitive on the global market. These policies were sustainable for quite some time because of continued investment and aid from developed countries, but once money became scarce, as happened with the 1973 Organization of Arab Petroleum Exporting Countries oil embargo in response to U.S. support of Israel during the Yom Kippur War and the subsequent 1979 reduction in oil supply due to the Iranian Revolution, investment and aid tended to become scarce as well. At this point, countries practicing ISI, some for 20 or 30 years, were faced with an increased need to borrow money; however, when interest rates rose, repayment became more difficult, further adversely affecting their economies. When Mexico declared that it could not meet its debt repayment obligations in 1982, lending virtually stopped and developing countries found themselves having to reconsider their ISI policies (Geddes, 2002).

Given the long-term effects of ISI policies and the results of the empirical research described previously, developing countries have had to reconsider their position on foreign investment, aid, and trade. Since the initiation of the Doha Round of World Trade Organization (WTO) talks in 2001, developing countries have increasingly attempted to renegotiate the terms of trade and demand concessions and exceptions to various provisions of the WTO they feel are discriminatory or that maintain their structural disadvantage. We also see developing countries working together, acting in loose blocs, both within the WTO and in negotiations with the European Union, to increase their leverage.

Future Directions

Although dependency and world-systems theory have significant flaws and have not been consistently verified by empirical research and observable development patterns, the underlying concepts and arguments are still appealing to many in the (semi-) periphery and the core. As dependency theory has evolved, theorists have had to accept the limitations of the theory as originally formulated and have instead begun to focus on more specific issue areas. Examples of

more narrowly focused research that is likely to prove fruitful for future research include investigation into the effects of globalization on dependence (see, for example, the special issue of *Studies in Comparative International Development* edited by Heller, Rueschemeyer, & Snyder, 2009), how dependence affects gender issues within the periphery (Scott, 1995), and dependence and political violence (Boswell & Dixon, 1990).

Another potential direction for the dependency approach has its roots in the failure of both dependency and modernization theory to be consistently supported when subjected to rigorous empirical testing. Neither is able to explain development (or its failure) on its own, yet both appear to be at least partially relevant. As a result, there has been a call by some theorists to reconcile and combine the two theories in an attempt to create a true universal theory of development. However, since the two are opposites, or mirror images of each other, there is concern that attempts to reconcile them would create an overly general theory that lacks useful and relevant explanatory power. Nevertheless, attempts to combine them could provide useful insight into the theories, with the potential to spark interesting new approaches to dependency.

Conclusion

Dependency and world-systems theory emerged in the 1950s and 1970s to refute the predominant approach to political and economic development (modernization theory) and provide an alternative explanation as to why developing countries fare so poorly. Their approach was to focus primarily on a singular external influence, the international capitalist economic system, and how a country's insertion into this system affects its ability to generate economic growth and development. Countries that enter the system as a result of decolonization or other forms of economic exploitation, or that entered well after the system's creation, are at a distinct disadvantage since the system is designed to benefit the developed industrialized economies of the core and to maintain their economic supremacy. Dependency theory, as a result, argues that although direct exploitation via colonialism has ended, dependence on the core is maintained via the economic control created by neocolonialism. Although dependent development modified dependency to argue that developing countries are no longer consigned to being helpless victims of exploitation by the core, they are still heavily impacted by it, and although some countries may be able to significantly improve their conditions and foster economic growth and development, it will not be enough to reach parity with the core. For those countries unable to attract foreign investment and loans to facilitate industrialization, the prospects remain grim.

This continuing focus on predominantly external factors continues to be a major weakness of dependency theory

since it virtually ignores internal factors such as social problems, ethnic conflict, and ineffective political institutions. Dependency theorists critiqued modernization theory for focusing exclusively on internal factors influencing development yet ultimately made the same assumption themselves by focusing exclusively on external influences—and in so doing left the dependency approach vulnerable to the criticism of being as biased and incomplete as modernization theory. Given the variety of experiences that have shaped developing countries, it is problematic to argue that regardless of colonial status, region, or domestic history, all are influenced in the exact same way by a limited set of influences. Because of these criticisms and the failure to establish clear and consistent evidence for the theory's assumptions and implications through empirical testing, mainstream dependency theory has moderated the breadth of its claims with recent scholarship, focusing instead on more specific issue areas rather than continuing to develop a universal theory with only limited support.

References and Further Readings

- Amin, S. (1974). *Accumulation on a world scale: A critique of the theory of underdevelopment*. New York: Monthly Review Press.
- Amin, S. (1977). *Imperialism and unequal development*. New York: Monthly Review Press.
- Amin, S., Arrighi, G., Frank, A. G., & Wallerstein, I. (1982). *Dynamics of global crisis*. New York: Monthly Review Press.
- Baran, P. (1957). *The political economy of growth*. New York: Monthly Review Press.
- Bornschier, V., Chase Dunn, C., & Robinson, R. (1978). Cross national evidence of the effects of foreign investment and aid on economic growth and inequality: A survey of findings and a re analysis. *American Journal of Sociology*, 84(3), 651–683.
- Boswell, T., & Dixon, W. J. (1990). Dependency and rebellion: A cross national analysis. *American Sociological Review*, 55(4), 540–559.
- Bradshaw, Y. W. (1985). Dependent development in black Africa: A cross national study. *American Sociological Review*, 50(2), 195–207.
- Cardoso, F. H. (1973). Associated dependent development: Theoretical and practical implications. In A. Stepan (Ed.), *Authoritarian Brazil* (pp. 142–178). New Haven, CT: Yale University Press.
- Cardoso, F. H., & Faletto, E. (1979). *Dependency and development in Latin America*. Berkeley: University of California Press.
- Chase Dunn, C. K. (1975). The effects of international economic dependence on development and inequality: A cross national study. *American Sociological Review*, 40(6), 720–738.
- Dos Santos, T. (1970). The structure of dependence. *American Economic Review*, 60(2), 231–236.
- Evans, P. (1979). *Dependent development: The alliance of multi national, state, and local capital*. Princeton, NJ: Princeton University Press.
- Frank, A. G. (1966). The development of underdevelopment. *Monthly Review*, 18, 17–31.
- Frank, A. G. (1967). *Capitalism and underdevelopment in Latin America: Historical case studies of Chile and Brazil*. New York: Monthly Review Press.
- Frank, A. G. (1978). *World accumulation, 1492–1789*. New York: Monthly Review Press.
- Furtado, C. (1967). *Development and underdevelopment*. Berkeley: University of California Press.
- Geddes, B. (2002). The great transformation in the study of politics in developing countries. In I. Katznelson & H. V. Milner (Eds.), *Political science: The state of the discipline* (pp. 342–370). New York: W. W. Norton.
- Heller, P., Rueschemeyer, D., & Snyder, R. (Eds.). (2009). Dependency and development in a globalized world [Entire issue]. *Studies in Comparative International Development*, 44(4), 287–456.
- Jackman, R. W. (1982). Dependence on foreign investment and economic growth in the third world. *World Politics*, 34(2), 175–196.
- Kaufman, R. R., Chernotsky, H. I., & Geller, D. S. (1975). A preliminary test of the theory of dependency. *Comparative Politics*, 7(3), 303–330.
- Marshall, S. E. (1985). Development, dependence, and gender inequality in the third world. *International Studies Quarterly*, 29(2), 217–240.
- Marx, K., & Engels, F. (2002). *The communist manifesto* (G. Stedman Jones, Ed.). London: Penguin.
- McGowan, P. T., & Smith, D. L. (1978). Economic dependency in black Africa: An analysis of competing theories. *International Organization*, 32(1), 179–235.
- O'Donnell, G. A. (1973). *Modernization and bureaucratic authoritarianism*. Berkeley: University of California Berkeley, Institute of International Studies.
- Prébis, R. (1950). *The economic development of Latin America and its principal problems*. New York: United Nations.
- Ray, J. L., & Webster, T. (1978). Dependency and economic growth in Latin America. *International Studies Quarterly*, 22(3), 409–434.
- Scott, C. (1995). *Gender and development: Rethinking modernization and dependency theory*. Boulder, CO: Lynne Rienner.
- Smith, T. (1979). The underdevelopment of development literature: The case of dependency theory. *World Politics*, 31(2), 247–288.
- Sunkel, O., & Fuenzalida, E. (1979). Transnational capitalism and national development. In J. Villamil (Ed.), *Transnational capital and national development: New perspectives on dependence* (pp. 67–93). London: Harvester Press.
- Wallerstein, I. (1974). *The modern world system: Capitalist agriculture and the origins of the European world economy in the sixteenth century*. New York: Academic Press.
- Wallerstein, I. (2004). *World system analysis: An introduction*. Durham, N C: Duke University Press.

40

FOREIGN POLICY ANALYSIS

BARIS KESGIN

University of Kansas

The scholarly study of foreign policy in the field of international relations (IR) goes back to the mid-1950s and early 1960s. Although this means that foreign policy analysis (FPA) is still a relatively young area of specialization, the FPA literature is rich and diverse. As it is discussed in the following pages, since its inception FPA has evolved through distinct stages (Hudson, 2005; Neack, Hey, & Haney, 1995). Specifically, FPA has moved away from searching for an overarching theory of foreign policy to finding theories that work under certain conditions. In this chapter, FPA refers to the scholarly study of foreign policy as a whole, which includes such distinct literatures as comparative foreign policy and foreign policy decision making.

This chapter proceeds as follows: First, foreign policy is described, and then the FPA subfield is situated within the broader field of IR. This is followed by a review of major approaches within this subfield from foundational works to the most contemporary research. At the end, the policy relevance of the work produced by the FPA scholars and future directions of the subfield are discussed. The chapter concludes with a list of suggested readings.

Introduction

What is foreign policy, and how is it different from the study of IR more generally? By *foreign policy* we mean

the actions, strategies, and decisions directed at actors outside the borders of a domestic political system (i.e., a state). That the primary intended target of policy is external to the domestic sphere distinguishes foreign from domestic policy (Kaarbo, Lantis, & Beasley, 2002). In other words, as Breuning (2007) puts it, foreign policy is “the totality of a country’s policies toward and interactions with the environment beyond its borders” (p. 5). A state’s foreign policy covers a variety of issues ranging from the rather traditional security and economic areas to environmental and energy issues, foreign aid, migration, and human rights (Breuning, 2007). The actors that initiate foreign policy actions, and those who are the targets of the actions, are often states—but not always.

According to Breuning (2007), FPA is “first and foremost interested in explaining how and why . . . decisions came about” (p. 164). In effect, this means that—distinct from the broader study of IR—FPA places an emphasis on the human beings involved in the decision-making processes (Breuning, 2007; Hudson, 2005, 2007). As such, FPA looks beyond the state as a single entity and includes the study of multiple actors within a state, both as individuals and as groups of individuals in their capacity to make or influence foreign policy. In other words, this inquiry goes further than governments and investigates the influence of individual leaders, bureaucracies, and institutions in foreign policy making.

FPA is an eclectic subfield of IR and benefits from many other disciplines such as psychology. One of the most important features of FPA is that it is pursued in a comparative perspective. Comparison entails searching for potentially generalizable patterns in foreign policy making across time, space, and issues. As such, this inquiry can involve comparing foreign policy choices of different leaders within the same country, or foreign policy decisions of different countries regarding the same or similar issues. To the degree that such work would empirically confirm theoretical propositions, a generalizable theory is reached—often with certain set of conditions attached to it (called a *midlevel* theory). As major works in this research are reviewed in the following sections, this chapter outlines the evolution of FPA as a subfield of IR research.

Theory

There is no single, distinct theory of foreign policy; instead, theories of foreign policy derive either from IR theories such as realism and liberalism or from various approaches to explaining domestic sources of foreign policy making within a state—for instance, leaders, bureaucracies, and culture. This difference in emphasis, respectively, corresponds to looking at external and systemic factors and societal and domestic sources in foreign policy. The latter make up the bulk of the literature in FPA. In contrast to the study of IR more generally, where there is the motivation to formulate grandiose, comprehensive theories, FPA is occupied with building midlevel theories. This chapter first briefly reviews IR theories that are relevant to explanations of and prescriptions for foreign policy; then, major approaches to domestic sources of foreign policy are discussed.

Theories of International Relations

IR theories explain how states relate with each other in world politics. A logical conclusion of this focus is explanations of foreign policy behavior in many IR theories (Kaarbo et al., 2002). Indeed, as Smith (1987) argues, “Virtually every attempt to explain international relations involves an explanation of foreign policy” (p. 348). Many, if not all, IR theories concentrate on the impact of the international system on foreign policy. Although all the theories discussed in this section here are discussed in greater depth elsewhere in this volume, a short review will help to place foreign policy within the context of the field of IR.¹

Realism

Realist theories of international relations are defined by their assumptions of anarchy and self-help, and a unified, rational actor understanding of the state. Accordingly, realists claim that in order to survive, states should act as power

maximizers. For realists, a state’s foreign policy is very much conditioned by its position within the international system and the distribution of power therein.

Liberalism

Although certain variants of liberalism share some assumptions of realism, liberals differ from realists in that, for them, the international system by nature is more conducive to cooperation. According to liberal theories of IR, it is in the self-interest of states to cooperate with each other. International institutions play a crucial role in facilitating cooperation among states, since they help states overcome distrust via their established rules. In contrast to realism, liberalism recognizes that domestically states are home to diverse interests and actors. As such, liberal theories of international relations also look into domestic politics since they explain state behavior. For instance, the role of interest groups or businesses in foreign policy is included in the analysis. Liberalism’s most important conclusion for foreign policy is that with a shared culture of liberalism and its ramifications on domestic institutions, like-minded liberal states would have peaceful relations with each other—the democratic peace argument.

Constructivism

Constructivism can best be considered as an approach to IR than as a theory. Constructivist approaches ushered in a major turning point in the study of IR and hence have significant repercussions for the study of foreign policy. Broadly speaking, constructivism defends that intersubjectively created social norms and values explain the behavior of actors in the international system. As such, constructivists question the very existence of such concepts as anarchy and argue that these concepts reflect our own understandings of international relations.

Other Theories of International Relations

Today, variants of realism, liberalism, and constructivism are major IR theories. Alternatives include but are not limited to feminist and Marxist approaches to IR. Similar to the dominant theories discussed previously, the alternatives are not foreign policy theories either, but again we can infer from feminism and Marxism as we study foreign policy.

Feminist approaches to international relations focus on issues pertaining to gender relations and question how these affect the way international relations is studied and practiced. Such an approach often entails how the exclusion of women from politics (or the society or both) and dominant, contrasting notions of masculinity affect world politics.

A Marxist understanding of IR emphasizes the importance of class relations in world politics. By extension, foreign policy read through the lens of Marxism would

explain foreign policy decisions in terms of economic interests and conflicts within and between states.

Approaches to Foreign Policy Analysis

The emergence of FPA as a distinct area of study dates back to the seminal Snyder, Bruck, and Sapin (2002, originally published in 1954) and two other works that followed (discussed in the next section). Since then, the field has significantly expanded, unpacking the domestic determinants of state behavior. Today, many scholars distinguish two phases in the development of the literature on FPA. The first stage includes the foundational works in the field; the second follows with a proliferation of research in psychological analysis of individuals and groups as foreign policy decision makers. This chapter follows the same approach in its discussion of the FPA literature.

Foundational Texts

FPA within the IR field goes back to three “paradigmatic works” written in the 1950s and 1960s (Hudson, 2005, 2007): *Decision Making as an Approach to the Study of International Politics* by Richard Snyder, H. W. Bruck, and Burton Sapin (2002); *The Ecological Perspective on Human Affairs With Special Reference to International Politics* by Harold Sprout and Margaret Sprout (1965); and “Pre-Theories and Theories of Foreign Policy” by James Rosenau (1966).

In reaction to the dominant realist approach to IR, in 1954, Snyder, Bruck, and Sapin (2002) call for looking beyond the state and specifically to actors involved in foreign policy decision making. Likewise, Sprout and Sprout (1965) suggest that the psycho-milieu of the individuals and groups involved in foreign policy making need to be understood in order to explain foreign policy behavior. Finally, Rosenau’s (1966) pretheories article, on the other hand, is a call for systematic, scientific, cross-national generalizations of state behavior. It was from this particular quest that the comparative foreign policy (CFP) approach was born, which sought to create a grand theory of foreign policy (one that would explain all foreign policy across time and space in empirical terms).

The significance of these works was that they championed the idea that individual decision makers and their characteristics were at the heart of understanding foreign policy (Hudson, 2007). Likewise, their message called for an understanding of foreign policy beyond a mere output but also as a process. Hence, the study of FPA has expanded much beyond and benefited from many other disciplines as it grew as a field of research.

Contemporary Approaches to Foreign Policy Analysis

Contemporary research and interest in FPA has not progressed in a linear fashion. First of all, one observes a

break with the Rosenau (1966) line of research that was engaged with producing a grand theory of foreign policy. By the early 1970s, such attempts at creating large data sets as the Comparative Research on the Events of Nations (CREON) and the Dimensionality of Nations (DON) were being overshadowed by an emerging literature focusing instead on explaining the impact of individual and group behavior on foreign policy. The CFP approach that derived from Rosenau and dominated the field during the 1970s gradually disappeared in the late 1980s, since it did not produce the grand theories of foreign policy it aspired to. In the meantime, use of various methodologies and disagreements in how to study foreign policy pushed the subfield in a short-lived crisis from the 1980s into the early 1990s. One of the reasons behind this temporary crisis was the transition from the search for a grand theory and solely empirically oriented research (the gradual vaporization, or more accurately reformulation, of CFP research) to the emergence of multiple approaches and methodologies. In addition, the unfolding events of the late 1980s that led to the end of the cold war and hence led to a new world of international politics also affected the study of foreign policy.

By the early 1990s, FPA as a field emerged with a relative consensus that midlevel theories served as new goals rather than a grand theory of foreign policy. As such, these theories explicitly stated their limitations and the conditions of applicability. In this process, the CFP approach also refined its search for a grand theory of policy as it looked for new directions in the study of foreign policy.² Since then, midlevel theories have shaped the FPA literature.

Contemporary approaches to the scholarly study of foreign policy have developed within this context and can be divided into literatures that focus on (a) individual, (b) group, and (c) societal characteristics of foreign policy actors. In the following section, first each is discussed as to their assumptions about the importance of these levels of analysis; then, their applications to foreign policy cases and empirical evidence for each are discussed.

Applications and Empirical Evidence

Theories of International Relations

Direct applications of IR as theories of foreign policy remain few in number. Despite the rarity of such work, we continue to benefit from the grand theoretical lenses of IR theories. For instance, a rare exception is Rittberger’s (2001) and his colleagues’ attempt to explain Germany’s foreign policy since its unification. A somewhat similar work is Sterling-Folker’s (2006) *Making Sense of International Relations Theory*. Otherwise, some realists attempt to apply realist theory to foreign policy (see Wivel, 2005); however, some realists debate this. In addition, Houghton (2007) discusses the gains from cooperation

between constructivism and FPA. These notwithstanding, the majority of IR research explains state behavior as an outcome of overarching processes or the international system.

Approaches to Foreign Policy Analysis

Arguably, the most important contributions to the sub-field of FPA came from domestic-based explanations. Responding to a call to understand the foreign policy decision-making process and benefiting from various fields of research, many works contributed to the expansion of knowledge in the study of foreign policy.

Individuals

Foreign policy decision makers are individual actors, and as all individuals, they too are bounded by the human mind. Hence, individual characteristics of decision makers (such as their beliefs, experiences, emotions, and conceptions of the self and nation) can have significant impact on foreign policy decision making. This effect may be higher under certain circumstances, such as during crises when an individual has to make a decision under stress and time pressure and possibly with limited information.

Research on individual characteristics of leaders benefited significantly from the field of psychology. For instance, in *Perception and Misperception in International Politics*, Jervis (1976) illustrated the importance of psychological factors in his investigation of the impact of misperceptions in foreign policy. Likewise, Vertzberger's (1990) *The World in Their Minds* looked at the impact of information processing on foreign policy decision making. Various other works deriving from psychological approaches all share the assumption that studying political leadership offered a lot to explain foreign policy. The interest in the psychological characteristics of decision makers has expanded to various specific topics; here, operational code and leadership trait analysis are discussed as examples. Other relevant research studies the impact of bounded rationality, motivated and unmotivated bias, cognitive maps, scripts, and schemas on foreign policy making.

Operational code analysis entails the study of the core belief system of an individual leader. In most contemporary examples of operational code analysis, operational code refers to a leader's beliefs about the political world and also his or her approach to political action. Operational code research goes back to Nathan Leites's works on the Soviet Politburo and the Bolshevik Revolution. Later, George (1969) further developed the concept as he refined operational code into a belief system composed of five philosophical beliefs and five instrumental beliefs (also see Holsti, 1970; Walker, 1977). The most significant stride in this research was the introduction of automated content analysis in the late 1990s (for a recent review, see Schafer & Walker, 2006a). Since then, the less time-consuming nature of machine coding attracted more attention to operational

code analysis. Among others, for instance, Schafer and Walker (2006b) illustrated that Tony Blair and Bill Clinton had different views of democracies and nondemocracies, and their actions also differed accordingly. They show that democratic leaders hold a cooperative orientation toward other democracies.

Another important approach to how psychological characteristics of political leaders affect their foreign policy choices is the Leadership Trait Analysis (LTA) framework, a research identified with Margaret Hermann. This particular line of inquiry derives from the assumption that leaders have different styles of decision making because leaders interact with their subordinates, close advisers, or other leaders in different manners, and they follow a different set of principles or rules as they interact with others. Hermann's decades-long research now covers 122 national leaders and 87 heads of state from around the world and suggests a set of different leadership styles. According to Hermann (2003), the most useful traits in assessing leadership style are (a) the belief that one can influence or control what happens, (b) the need for power and influence, (c) conceptual complexity (the ability to differentiate things and people in one's environment), (d) self-confidence, (e) the tendency to focus on problem solving and accomplishing something versus maintenance of the group and dealing with others' ideas and sensitivities, (f) general distrust or suspiciousness of others, and (g) the intensity with which a person holds an in-group bias. The LTA has proved to be a fruitful line of research (among many others, Kaarbo & Hermann, 1998). Recently, Dyson (2006) showed the significance of Tony Blair's personality and leadership style in explaining British foreign policy making during the 2003 Iraq War. Dyson found that Blair had a high belief in his ability to control events, a low conceptual complexity, and a high need for power; Blair's preferences and behavior as derived from his personality profile corresponded to Blair's Iraq decisions. Dyson's work is a good illustration of how individuals matter in foreign policy decision making.

Groups

Individual decision makers often interact with other individuals as foreign policy decisions are formulated. These interactions may take place in either small or large groups but always under the unique settings of the group that are created purposefully or inadvertently. Hence, the structure of a group of individuals and the process of decision making therein become a concern for scholars of foreign policy. For instance, Kaarbo (1998) studied the dynamics of coalition cabinet decision making in parliamentary systems (also see Ozkececi-Taner, 2005). As discussed later, this concern mainly derives from the fact that individuals in group settings strive to conform to others—which hampers decision making in a way that reduces alternatives being discussed or leads members of the group to wrongful interpretations of the reality.

The study of decision-making dynamics in small and large groups also benefited from the field of psychology and has continually attracted foreign policy scholars to conduct further research about its effect on foreign policy decision making. The groundbreaking work was Janis's (1972) *Victims of Groupthink*. In this book, Janis illustrated that small groups of decision makers were prone to ignore, misinterpret, or even reject new information in order to avoid controversy and lack of cohesion within the group. During the 1990s, further research on this topic explored what happens beyond groupthink (Hart, Stern, & Sundelius, 1997), suggesting that there is more to explore than merely assuming that groups act differently and that contextual factors may affect the decision-making processes in groups.

A similar line of inquiry looked at the conflicts between different leaders and groups involved in decision-making processes. Initially introduced by Allison (1969), the bureaucratic politics approach, one of the three conceptual models used to explain the Cuban missile crisis, generated a lot of attention in comparative foreign policy analysis.³ Indeed, Allison's work about decision making in the U.S. and Soviet administrations during the crisis most deservedly became "the best known example of bureaucratic politics" (Kaarbo, 1998, p. 69). According to the bureaucratic politics model (BPM), policymakers compete with each other in a game of bargaining within the hierarchy of government. The model suggests that the hierarchical nature of decision making and power sharing promotes differences among players (Allison, 1969). Because every actor has different sets of objectives, actions result from bargaining games among them; as such, the BPM describes them as resultants rather than outputs. That is, the outcomes of bargaining games are often different from what was sought by each actor. Allison illustrates this approach with an analogy of foreign policy behavior with moves in the game of chess where "a number of distinct players, with distinct objectives but shared power over the pieces, were determining the moves as the resultant of collegial bargaining" (p. 691). Although the model faced harsh criticisms and was even discredited (Bendor & Hammond, 1992), one of the reasons for the continued interest in the BPM is the availability of new information about the Cuban missile crisis. As the official records of the Cuban missile crisis have been declassified, scholars of foreign policy have also refined the BPM.

Recent developments too have led to a renewed interest in the BPM. Specifically, policymaking after the September 11, 2001, attacks and in the months leading to and following the 2003 Iraq War during the administration of George W. Bush has attracted a lot of attention. Several studies discussed the competitive and fractionalized nature of decision making within the U.S. administration. It is, however, interesting to note that the BPM gains impetus from decision-making processes at times of serious conflict (Cuban missile crisis) or war (Vietnam and Iraq Wars; see Hudson, 2005).

In her research on minority influence over foreign policy decisions of coalition governments, Kaarbo (1998) argued for the merits of the perspective (power sharing, competitive decision making) and suggested that the model (its predictions and assumptions) itself might be dropped. Overall, along with these, possibly one last reason for the BPM's presence in the study of foreign policy has been that it is a simple, plausible, and logical framework (Rhodes, 1994).

Society

As factors beyond the individual and group level, national characteristics of a state can have a significant effect on its decision-making processes. The historical, socioeconomic, and political context of a society can all exert influence in foreign policy making. Societal sources of foreign policy range from culture and identity to elite and public opinion.

Now relatively an extinct research at this level of interest, but a lively one during the 1970s (see, e.g., East, 1973), is the effect of national attributes (size, political, and economic system, etc.) on foreign policy (where engaging in a war was often, if not always, the foreign policy event).⁴ Presently, the democratic peace research can be considered a remaining example of such an interest in national attributes. In addition, a new generation of event data sets, such as the Kansas Events Data System (KEDS), also belongs to this tradition (Breuning, 2007; Hudson, 2005, 2007). Distinct from the past endeavors, these new data sets do not reject the vital role individual decision-makers' play in foreign policy decision making (Breuning, 2007).

The study of culture as it relates to foreign policy analysis is a societal level of inquiry as well. Culture, according to Kaarbo et al. (2002), can "set [broad] parameters for foreign policy" (p. 15). For instance, cultural characteristics of a society can influence its decision-making processes: As Kaarbo et al. (2002) discuss, if consensual decision making is the practice in a country, then the making of a decision can take longer. Despite such recognition, research about the influence of culture on foreign policy remains a relatively less studied issue area. A major exception is an edited volume, *Culture and Foreign Policy* (Hudson, 1997). In a truly comparative framework, this book investigated the influence of culture on foreign policy in diverse settings like Belgium and the Netherlands, China, and India. For instance, Chafetz, Abramson, and Grillot (1997) used national role conceptions as a conceptual tool to explain how Belarus and Ukraine developed different approaches to nuclear nonproliferation. Chafetz et al. argued that the possession of nuclear weapons was incompatible with the Belarussian national role conception and hence Belarus immediately signed the Nuclear Non-Proliferation Treaty (NPT); Ukraine, on the other hand, desired such possession and because of inconsistent, multiple national role conceptions moved between complying

with and defying the NPT. Barnett (1999) also looked at culture as a factor affecting foreign policy. According to Barnett, Israel's acceptance of the Oslo Accords can be explained by Israeli Prime Minister Rabin's creation of a political culture that presented withdrawal from the occupied territories as a desirable and legitimate option for Israel.

In contrast to the research about the impact of culture on foreign policy, the study of elite and public opinion in relation to foreign policy is a rich area of research. Questions pertaining to the effect of public opinion on foreign policy have been contested in the foreign policy literature for decades and are but settled; moreover, the answers are significantly changed since the initial studies on this subject. As Holsti (1992) explained, the conventional view of the role of public opinion in the United States (the Lippmann-Almond consensus) was that it was "ill-informed, emotion-driven, shortsighted, and self-absorbed" (p. 514). Although this initial conception of the relationship between public opinion and foreign policy has long been challenged since the 1970s (see Shapiro & Page, 1988), there is not a new consensus yet (Holsti, 2002). We know that this relationship is complicated by the fact that government officials too attempt to shape public opinion (Foyle, 2004). There is more interaction and competition between the government and the public in an effort to shape the former's policies and the latter's opinion. Last, the subject definitely demands more attention especially at another time of war. As Holsti (1992) argued, research on the role of public opinion in U.S. foreign policy has been shaped by three 20th-century wars: World Wars I and II and the Vietnam War. It is later argued that in the post-September 11 world, the Iraq War constitutes a special place in understanding public opinion and foreign policy. Once again, another war led to tensions between the policymakers and the publics in various parts of the world.

Policy Implications

To what extent does the scholarly work in the FPA subfield affect real world policy? One of the goals of FPA scholarship is to produce useful knowledge for the practice of foreign policy making. It is, however, safe to argue that the relationship between the scholarly work and the practice of policymakers has been tumultuous at best. Although academics at times complained about the nature of policies, policymakers often distanced themselves from the abstract world of theories. George (1993) argued that the main distinction between the communities of scholars and policymakers is that each has different professional goals and indeed they cannot easily communicate with each other.

That is not to say, however, that academics and policymakers do not engage with each other at all. To the contrary, every now and then individuals go back and forth between

academia and the policy world. For instance, in the United States, some notable examples are Henry Kissinger and Condoleezza Rice; other examples around the world are relatively less known—if there are any at all.⁵ According to George (1993), these are indeed the individuals who can contribute to bridging the gap between the scholarly world and the policymaking world.

Future Directions

One can argue that although FPA as a field of study steadily progressed, it was a half-empty glass with a capability to deliver more.⁶ That metaphorical analogy has a legitimate reasoning; indeed, there is a lot more to be explored. For instance, there is not much work yet on integrating different factors influencing foreign policy into a comprehensive framework, or in particular areas of interest such as public opinion, we are relying on some specific contexts—in that example, on the U.S. context. In addition to such gaps in the extant FPA literature, other developments such as the availability of automated research techniques (see below) make new approaches to FPA possible. Here, some of the gaps and new opportunities in FPA research are discussed as possible future avenues of research in FPA.

In preceding sections, contemporary approaches to the scholarly study of foreign policy—other than IR theories—are discussed by dividing the literature to individual, group, and societal factors in foreign policy. Notably, there is no mention of studies that integrate multiple factors or levels of analysis. Indeed, there is a lack of multilevel (integrative) work in FPA. A major exception to this statement is Putnam (1988), where he argues that decision makers often find themselves in a simultaneous two-level game between their domestic constituents and international pressures. Further research on two-level games, or multilevel research in foreign policy analysis, remains an important but not much developed research agenda.

As discussed earlier, there is often an increased interest in foreign policy decisions and decision-making processes during war times. In the FPA literature, bureaucratic politics of decision making and the public's role in foreign policy often benefit from this interest. Yet the changing circumstances of world politics create new challenges and demand more work in explaining foreign policy decisions during wars. For instance, the processes of decision making in the U.S. administration during the 2003 Iraq War call for a renewed attention to bureaucratic politics of foreign policy. Similarly, the somewhat coordinated nature of protests across the globe during initial months of the Iraq War (likewise during the World Trade Organization meetings) constitutes an interesting topic to investigate. Furthermore, in the extant literature about the role of public opinion on foreign policy, there is a heavy focus on the United States; FPA will definitely benefit from expanding this view and

including more comparative cases about the influence of public opinion on foreign policy. Specifically, there is a dearth of literature on the role of public opinion in non-democracies. As Telhami (1993) showed the conventional wisdom here may not necessarily be true, and FPA definitely needs more research on the role of public opinion on foreign policy in nondemocracies.

Last, as more data become available, quantitative and automated approaches to foreign policy analysis will increase both in quality and quantity. For instance, Kansas Events Data System (KEDS) illustrates a successful research program in this particular type of research. Additionally, renewed interest in automated at-a-distance methods such as ProfilerPlus, which enables computer-based leadership traits and operational code analysis, signals such a direction in FPA (see Schafer & Walker, 2006a).

Conclusion

FPA, a subfield of IR, has significantly contributed to our understanding of world politics since its inception in 1954 (Snyder et al., 2002). It has provided a rich menu of approaches to explain and understand foreign policy in a generalizable way and beyond system-oriented explanations. Although this chapter draws only a broad picture of this literature, as discussed FPA captures multiple levels of analysis with detailed investigations of the role of individuals, groups, and states or societies. FPA also stands as a successful example of interdisciplinary work, particularly for its collaboration with psychology in explaining the influence of the individual on foreign policy. Today, students, scholars, and policymakers have many diverse tools for explaining foreign policy thanks to those achievements.

Notes

1. The discussion about IR theories is generalized and does not distinguish between the many different variances within each theory.
2. One must look at the product of this search for a detailed discussion: Hermann (1987).
3. For an expanded version of Allison's (1969) article, see Allison (1999).
4. Event data are records of interactions among states in the open press, and illustrate these in a sequence of a numerical score or a categorical code (for a detailed discussion see Phil Schrodt's essay in Neack, Hey, & Haney, 1995).
5. Kissinger was a faculty member at Harvard University and served first as national security advisor and then as secretary of state for Richard Nixon. Rice was a faculty member at Stanford University before becoming national security advisor for George W. Bush, and then she served as secretary of state in Bush's second term after her service, Rice returned to Stanford. An example from outside the United States is Turkey's minister of foreign

affairs since May 2009, Ahmet Davutoglu, who was senior foreign policy adviser to the prime minister from January 2003 until May 2009.

6. See, for instance, Vertzberger (2002).

References and Further Readings

- Allison, G. (1969). Conceptual models and the Cuban missile crisis. *American Political Science Review*, 63(3), 689–718.
- Allison, G., & Zelikow, P. (1999). *Essence of decision: Explaining the Cuban missile crisis* (2nd ed.). New York: Addison Wesley.
- Barnett, M. (1999). Culture, strategy and foreign policy change: Israel's road to Oslo. *European Journal of International Relations*, 5(1), 5–36.
- Bendor, J., & Hammond, T.H. (1992). Rethinking Allison's models. *American Political Science Review*, 86(2), 301–322.
- Breuning, M. (2007). *Foreign policy analysis*. New York: Palgrave.
- Chafetz, G., Abramson, H., & Grillot, S. (1997). Culture and national role conceptions: Belarussian and Ukrainian compliance with the nuclear nonproliferation regime. In V. M. Hudson (Ed.), *Culture and foreign policy* (pp. 169–200). Boulder, CO: Lynne Rienner.
- Dyson, S. (2006). Personality and foreign policy: Tony Blair's Iraq decisions. *Foreign Policy Analysis*, 2(3), 289–306.
- East, M. (1973). Size and foreign policy behavior. *World Politics*, 25(4), 556–576.
- Foyle, D. (2004). Leading the public to war? The influence of American public opinion on the Bush administration's decision to go to war in Iraq. *International Journal of Public Opinion Research*, 16(3), 269–294.
- George, A. (1969). The "operational code": A neglected approach to the study of political leaders and decision making. *International Studies Quarterly*, 13(2), 190–222.
- George, A. (1993). *Bridging the gap: Theory and practice in foreign policy*. Washington, DC: United States Institute of Peace.
- Hart, P., Stern, E. K., & Sundelius, B. (Eds.). (1997). *Beyond groupthink: Political group dynamics foreign policy making*. Ann Arbor: University of Michigan Press.
- Hermann, M. G. (2003). Assessing leadership style: Traits analysis. In J. Post (Ed.), *The psychological assessment of political leaders* (pp. 178–212). Ann Arbor: University of Michigan Press.
- Hermann, C., Kegley, C. W., & Rosenau, J. N. (Eds.). (1987). *New directions in the study of foreign policy*. Boston: Allen & Unwin.
- Hill, C. (2003). *The changing politics of foreign policy*. New York: Palgrave.
- Holsti, O. (1970). The operational code approach to the study of political leaders: John Foster Dulles' philosophical and instrumental beliefs. *Canadian Journal of Political Science*, 3(1), 123–157.
- Holsti, O. (1992). Public opinion and foreign policy: Challenges to the Almond Lippmann consensus. *International Studies Quarterly*, 36(4), 439–466.
- Holsti, O. (2002). Public opinion and foreign policy analysis: Where we were, are and should strive to be. In M. Brecher & F. Harvey (Eds.), *Millennial reflections on international studies* (pp. 514–528). Ann Arbor: Michigan University Press.

- Houghton, D. P. (2007). Reinvigorating the study of foreign policy decision making: Toward a constructivist approach. *Foreign Policy Analysis*, 3(1), 24–45.
- Hudson, V. (2005). Foreign policy analysis: Actor specific theory and the ground of international relations. *Foreign Policy Analysis*, 1(1), 1–30.
- Hudson, V. (2007). *Foreign policy analysis*. Lanham, MD: Rowman & Littlefield.
- Hudson, V. M. (Ed.). (1997). *Culture and foreign policy*. Boulder, CO: Lynne Rienner.
- Janis, I. (1972). *Victims of groupthink*. Boston: Houghton Mifflin.
- Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.
- Kaarbo, J. (1998). Power politics in foreign policy: The influence of bureaucratic minorities. *European Journal of International Relations*, 4(1), 67–97.
- Kaarbo, J., & Hermann, M. (1998). Leadership styles of prime ministers: How individual differences affect the foreign policymaking process. *Leadership Quarterly*, 9(3), 131–152.
- Kaarbo, J., Lantis, J. S., & Beasley, R. K. (2002). The analysis of contemporary comparative foreign policies. In R. K. Beasley, J. Kaarbo, J. S. Lantis, & M. T. Snarr (Eds.), *Contemporary comparative foreign policy* (pp. 1–23). Washington, DC: CQ Press.
- Neack, L., Hey, J. A. K., & Haney, P. J. (Eds.). (1995). *Foreign policy analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Ozkececi Taner, B. (2005). The impact of institutionalized ideas in coalition foreign policy making: Turkey as an example, 1991–2002. *Foreign Policy Analysis*, 1(3), 249–278.
- Putnam, R. D. (1988). Diplomacy and domestic politics: The logic of two level games. *International Organization*, 42(3), 427–460.
- Rhodes, E. (1994). Do bureaucratic politics matter?: Some disconfirming findings from the case of the U.S. Navy. *World Politics*, 47(1), 1–41.
- Rittberger V. (Ed.). (2001). *German foreign policy since unification*. Manchester, UK: Manchester University Press.
- Rosenau, J. N. (1966). Pre theories and theories of foreign policy. In R. B. Farrell (Ed.), *Approaches to comparative and international politics* (pp. 115–169). Chicago: Northwestern University Press.
- Schafer M., & Walker, S. G. (Eds.). (2006a). *Beliefs and leadership in world politics: Methods and applications of operational code analysis*. New York: Palgrave.
- Schafer, M., & Walker, S. G. (2006b). Democratic leaders and the democratic peace: The operational codes of Tony Blair and Bill Clinton. *International Studies Quarterly*, 50(3), 561–583.
- Shapiro, R., & Page, B. (1988). Foreign policy and the rational public. *Journal of Conflict Resolution*, 32(2), 175–190.
- Smith, S. (1987). Foreign policy analysis and international relations. *Millennium: Journal of International Studies*, 16(2), 345–348.
- Snyder, R. C., Bruck, H. W., & Sapin, B. (2002). *Foreign policy decision making revisited*. New York: Palgrave.
- Sprout, H., & Sprout, M. (1965). *The ecological perspective on human affairs with special reference to international politics*. Princeton, NJ: Princeton University Press.
- Sterling Folker, J. (Ed.). (2006). *Making sense of international relations theory*. Boulder, CO: Lynne Rienner.
- Telhami, S. (1993). Arab public opinion and the Gulf war. *Political Science Quarterly*, 108(3), 437–452.
- Vertzberger, Y. (1990). *The world in their minds*. Stanford, CA: Stanford University Press.
- Vertzberger, Y. (2002). Foreign policy analysis: Steady progress and a half empty glass. In M. Brecher & F. P. Harvey (Eds.), *Millennial reflections on international studies* (pp. 479–501). Ann Arbor: University of Michigan Press.
- Walker, S. G. (1977). The interface between beliefs and behavior: Henry Kissinger's operational code and the Vietnam War. *Journal of Conflict Resolution*, 21(1), 129–168.
- Wivel, A. (2005). Explaining why state X made a certain move last Tuesday: The promise and limitations of realist foreign policy analysis. *Journal of International Relations and Development*, 8(4), 335–380.

41

FEMINIST INTERNATIONAL RELATIONS

KAREN J. VOGEL

Hamline University

This chapter is intended as a starting point for helping students understand the multiple perspectives and diverse approaches of feminist international relations (hereafter referred to as “feminist IR”). Rethinking topics of mainstream international relations (IR) in new ways, feminist IR critiques, expands, and deepens theories and methodologies that explain how global politics affect and can be impacted by gender. Feminist IR begins with exposing the experiences of women in international politics, explores how gender is socially constructed and what those social constructions do at both the local and the international levels, and uses gender as an analytical category in understanding the interconnectedness of international economic and political phenomena.

In this chapter, an overview of the literature on feminist IR is presented. First, some definitions and the evolution and history of feminist IR as a subfield are discussed, including a summary of the theoretical and methodological debates between feminists and mainstream IR scholars. Next, various methods of research and themes of feminist IR research are briefly described. Themes include perspectives on understanding war, militarism, and security; views of the state and sovereignty; discussions of globalization, development, and international political economy; and finally, transnational feminism and human rights. Third, the usefulness of feminist IR and how this research contributes to improving individuals’ lives as well as practical implications for international organizations are discussed.

Next, future directions for feminist IR are presented. And finally, the chapter concludes with an overall summary and a list of sources or references where students can learn more about the topic.

Beginnings: The Evolution of Feminist IR Theory

A good place to begin is with a brief discussion of definitions of feminism and international relations. Multiple definitions of feminisms exist in academic literature (see Tickner, 2002), and space here does not allow for detailed discussions of each. However, to give a general sense of these terms, feminist perspectives can be understood as philosophical theories, political views, and analytical approaches that call for social justice and the equal treatment of women, illuminate the nature of socially constructed and institutionalized definitions of gender, and seek the dismantling of oppressive structures in social, political, and economic life. In short, the key goals of most feminist agendas are to understand, to challenge, and to change women’s subordinate roles to men, whether that be in the community, the state, or the international system as a whole. These goals can be promoted by individuals of any gender, groups of any size or composition, meeting face to face or on the web, or working through states or international entities (Ferree & Tripp, 2006).

For many years, traditional, mainstream IR was understood in terms of the study of states and their foreign policies. The prevailing theoretical framework was realism that carried with it several key assumptions about states' interests and behaviors in the international system. In recent years, mainstream (IR) has broadened and seen the development of multiple theoretical and methodological perspectives. As Tickner (2002, 2005) found, constructivism, for example, which also helped to open the space for the emergence of feminist IR, appeared in the 1990s. It called for the acknowledgment that international structures and activities are not just determined by material forces but also based on socially constructed and shared ideas. Identities and interests of transnational actors are constructed by these shared ideas and not just objective facts of nature. Constructivism, as well as many other theoretical developments, expanded the understanding of IR as a field of study. Mainstream IR more generally is understood now as the study of global issues and their formation as well as the study of transnational actors, including states, intergovernmental organizations (IGOs), nongovernmental organizations (NGOs), and transnational corporations (TNCs) and their foreign relations.

Key themes in feminist IR theories and methodologies, therefore, combine elements of feminism and international relations. Feminist IR began with investigating women and their place in the world and analyzing why political science, and traditional IR in particular, seemed so distant from women's lives (Tickner, 2005). Over time, the sub-field has evolved to examining gender more broadly and exploring multiple aspects of inclusion and exclusion of gender in international political activities. In other words, feminist IR has become interested in more than just counting how few women served as ambassadors or in key leadership positions in international organizations. Feminist IR seeks to extend critical analysis by theorizing in ways that draw together race, class, gender, sexualities, and identities and by acknowledging that gendered structures exist in a variety of ways at local, national, and international levels.

Feminist IR tends to argue for understanding connections between both the domestic and international, starting at the bottom or with a microlevel perspective and moving up or to a more macrolevel understanding. The analysis of firsthand accounts, personal narratives, interviews, or participant-observer experiences in case studies tend to be more favored as important in shedding light on the complex interactions of gender and global and local phenomena. As many authors have observed, feminist IR takes the old feminist phrase "The personal is political" and transforms it to "The private is global" (Kantola, 2007), meaning that much can be learned about the world through this personal or face to face interaction.

Emergence of Feminist IR

Feminist theory and methodologies in general existed in other academic disciplines prior to impacting traditional

international relations and political science. For example, discussions about how humans know what they know, the need to be reflective and self-aware in one's research choices, and how to contextualize gender roles were prevalent in women's studies, sociology, philosophy, history, and literary studies in the late 1960s and 1970s. As Wibben (2004) recounted in her history of feminist IR, feminists were exploring male bias in research and methodology at interdisciplinary conferences as early as 1975—long before such questions entered the mainstream academic journals of political science or international studies.

However, as multiple perspectives of feminism developed, expanded, and permeated social science, academic work in political science, and international relations in particular, emerged as well. By the late 1980s and early 1990s, several conferences, journal articles, and books in political science featured empirical studies on women and global issues as well as discussions of various feminist theories and methods. Attention was being paid to identifying the use of gendered language in discourses on international issues; explaining the need for women to be recognized, not just as victims but as key players on the world stage; exposing the gendered nature of states and international organizations; and critiquing the male-dominated field of IR itself. Feminist IR argued that without understanding the relationships of gender and power in both domestic and international terms, traditional male-constructed categories of research and methodologies were missing half the picture. Only a partial view of the realities and inequalities facing men and women and in the political discourses of diplomacy, foreign policy, war, militarism, and security were being revealed.

Some early classic books dealing with these subjects include the following: *Bananas, Beaches and Bases* by Cynthia Enloe (2001, originally published in 1990), J. Ann Tickner's (1992) *Gender in International Relations: Feminist Perspectives in Obtaining Global Security*, and Peterson and Runyan's (2010) *Global Gender Issues*. More recent edited collections followed, focusing on debates about feminist realism and quantitative research versus postmodern, qualitative methodological approaches (See Ackerly, Stern, & True, 2006; Jonasdottir & Jones, 2009). Each of these works contributed in different ways to substantial, new insights and methods to study the role of gender in international politics.

Along with books and edited collections, specific academic conferences, panel presentations, and journal publications provided forums for the discussion of methodologies and theoretical directions. Since the 1990s, professional meetings of the International Studies Association, British Political Science Association, and American Political Science Association have all sponsored panels and special topic sections about feminist IR, including feminist IR epistemology; feminist theories and the state; feminist approaches to understanding war, militarism, and security; the emergence of transnational feminism; and the role

of gender in the policies of international organizations, just to name a few.

A new specialized, scholarly journal, the *International Feminist Journal of Politics*, founded by Jan Jindy Pettman of the Australian National University of Canberra in 1999, also legitimized feminist IR as an academic field of study. This peer-reviewed journal uses themed issues to address a number of feminist IR topics such as gender in conflict and post-conflict societies; ideologies, religions, and conflict; and by 2009, human rights and feminism. This journal, as well as the publication of numerous articles in other academic sources, such as the *British Journal of Politics and International Relations* (BJPIR), *Foreign Affairs*, *Gender and Politics*, and *International Studies Quarterly*, made visible feminist IR within the academic communities of political science and international studies.

Critiques and Debates:

Traditional IR Versus Feminist IR

The emergence of feminist IR has not been without criticism, and continuing debates over epistemologies, methods, and questions for research occur. Several articles trace, describe, and analyze the trajectory of these debates over the last few years. First is Cynthia Weber's (1994) piece titled "Good Girls, Little Girls, and Bad Girls." This work was penned in response to Robert Keohane's (1989) critique of feminist IR in which he argued that feminist analysis could only provide a limited insight into international relations. According to Keohane, feminist perspectives could help point to where women had been victims of patriarchy or left out of political processes but lacked the empirical strength of building testable hypotheses.

As Weber (1994), Sylvester (1994), later Tickner (1997), and Wibben (2004) note, Keohane (1989) missed one of the key points of feminist IR: that the object of study can be impacted by the observer and the observation itself, that concepts have contextual relativity and are flexible, and that traditional IR does not fully account for the dynamic, fluid boundaries of international relations and gender. Traditional empirical research and so-called testable hypotheses can have their own male biases that by their very nature may not include or take into account complex gender elements.

In particular, J. Ann Tickner's (1997) seminal article "You Just Don't Understand: Troubled Engagements Between Feminists and IR Theorists" explored in more detail why misunderstandings continued between feminists and traditional IR theorists. Her claim was that feminist IR was based on assumptions and epistemologies that were divergent from those that underpin the traditional discipline. Therefore, feminist IR did not fit easily with state-centric and structural, positivist approaches normally taken by traditional IR scholars. But that did not mean feminist IR should be ignored or dismissed as a lesser form of inquiry. Something was to be said for a menu of complex, rich, and diverse theoretical and methodological approaches, and feminist IR research with qualitative as well as quantitative

methods was needed to understand complex transnational, gendered relationships (see Tickner, 2005).

Following Keohane, Francis Fukuyama (1998) in his work on women and international politics pointed to the limitations of feminist IR. He suggested that feminists tended only to see men as aggressive and women as peaceful. As found by Wibben (2004), many feminist IR scholars responded to Fukuyama, arguing that he completely misread, oversimplified, or missed entirely the variety of theories and methodologies promoted by feminist IR. He did not understand that feminist scholarship revealed the fluid nature of definitions of gender and how the impacts and influences of those definitions played out politically. Feminist IR research did not assume that men and women always behaved the same way in every context but rather looked at the varieties of femininities and masculinities as they played out in gendered structures and institutions.

Feminist IR illuminated the male dominance in the field as a whole and affirmed the need to engage in feminist-informed research. As Charlotte Hooper (2001) noted in her work, *Manly States*, everything from the nature of the subjects to the methods of research to the lack of females in academic and international careers culminated in marginalizing and making invisible women's roles both as subjects of research and practitioners in the field. Feminist IR could expose gender-based distinctions in a number of traditional areas of IR research and explore a variety of research methods in the process, even if traditional IR still sought to marginalize the attempt. Sylvester (2002) further encouraged feminist researchers to think about the international and draw inspiration from other feminist writers. Interestingly enough, what has emerged in the last few years is a rich array of themes, methods, and knowledge that has provided additional insight into key aspects of international politics.

Themes in Feminist IR and Methods of Research

Just as debates have existed between traditional IR scholars and feminist IR researchers, differences in approaches, themes, and methods have also emerged with the feminist IR community. Feminist IR is by no means monolithic or wedded to one particular issue or topic within international studies. Indeed, many feminist IR scholars prefer not to be identified with one approach or a single, traditional disciplinary label (Tickner, 2005). As this section shows, various applications of feminist IR theories and methods abound. A brief description of types of methods used by feminist IR scholars are presented next, followed by a sample of some of the areas and questions explored in the feminist IR literature.

Types of Methods

Tickner is perhaps one of the most well-known and prolific feminist IR writers to discuss these issues. From her

earliest articles and books in the 1990s to her discussion of feminist IR methodological questions in the *International Studies Quarterly* (2005), she has provided key insights into the patterns of feminist IR work. She has described and analyzed many of the questions that feminist IR theorists ask as they engage in the process of research as well as the types of research tools they employ.

Tickner (2005) affirmed that no single or unique method of feminist IR exists and a variety of research tools are necessary to address gender and international politics. However, adapting research methods from other disciplines is one of the hallmarks of feminist IR. For example, ethnography, a tool more often used by anthropologists; literary theory and analysis of life narratives, borrowed from studies of literature and humanities; alternative ways of seeing or locating what is missing from a picture or situation, taken from the study of fine arts; and analyses of symbols, language, and phrasing, taken from communication studies and rhetoric, have all been used as interpretative methods of research. Case studies, philosophical argument, and participant-observer techniques too have been applied in various ways to provide insights into gender and its connections between local and global politics (see also Ackerly et al., 2006; Prügl, 2007).

Feminist scholars tend to view quantitative techniques with some skepticism and wariness when trying to talk about inequality or forms of subordination. Gender inequality, for example, may not be easily indexed or measured by a single variable because of the complex historical, sociocultural power relationships, or even differing perspectives regarding the understanding of what the term *inequality* means (Tickner, 2005). Waring (1988) has also elaborated on problems of doing quantitative work with large data sets that rely heavily on information gathered from state accounting systems. Often, governments do not include statistics on women's unpaid labor, for example, or clear information about who really is head of the household. Single women or women of a particular ethnicity or class may be completely left out of certain categories (Waring, 1988). Through the 1980s, even cross-national statistics on development or human rights supplied by the United Nations or the European Union were incomplete or inadequate when describing gender differences and therefore suspect if used in large quantitative studies to explain some gender patterns.

More recently, some feminist IR scholars have returned to quantitative techniques or found ways to combine both qualitative and quantitative research. Caprioli and Boyer (2001), using the International Crises Behavior Project data set, employed multinomial logistic regression to analyze whether a connection existed between domestic gender equality and states' use of violence internationally. With improvements in data collection in United Nations and European Union agencies since the late 1990s, as well as attempts at gender mainstreaming (bringing women into the decision-making process and taking into account the impact on women of policy) within these organizations, renewed

interest in using quantitative tools has emerged within some parts of the feminist IR community, especially when considering issues surrounding women's labor and the global economy or issues of human security. However, feminist IR scholars still caution that one must not be unreflective in thinking about how specific data are collected or for what purpose (Carney, 2004). How quantitative research can evolve and be made more applicable for feminist IR remains an ongoing question for exploration and debate.

Themes in Feminist IR

No matter what their methods, as Prügl (2007) points out, feminist IR scholars are remarkably reflective, self-aware, and conscious about their relationship to their subjects of study. They seek to make visible or to bring to light ideas, experiences, and phenomena related to gender and international politics that may have been hidden or ignored. This can be seen in their approaches to themes of war, militarism, and security; sovereignty and the state; globalization, development, and international political economy; and transnational feminism and human rights. Within each of these topics, feminist IR scholars attempt to show how an analysis of gender-based distinctions can enhance knowledge about human beings and the world. And even though these themes are set out in separate subsections here, quite often feminist IR scholars point to the interconnections and influences of each on the other.

Sovereignty and the State

One of the earliest areas of inquiry for feminist IR revolved around a critique of core IR concepts, in particular, understandings of the state and sovereignty. Kantola (2007) has provided an excellent summary and analysis of feminist theories and research about gender, the state, and issues of sovereignty. She identifies three general areas of literature and in the process unpacks concepts of gender, power, and state characteristics. The first category of literature incorporates both feminist IR and comparative politics perspectives and methods. The main goal of this work was to ask questions about the inferior position of women, the lack of position for women, or both in the social, economic, and political processes of the state. This literature revealed the continuing absence of women in decision-making roles related to foreign policy and international affairs, whether that be in terms of being elected to political office or appointed to key positions in key decision-making arenas (Kantola, 2007).

Women and the state are clearly defined variables. The state tends to be seen as a unitary entity, with identifiable institutions and structures. The actions of states can be detrimental or beneficial to women, and power is described in a top-down, binary fashion—either one has it or does not. The category of women also tends not to be broken down or disaggregated to take into account the different ethnicities or economic backgrounds of women.

What becomes important is simply to show where women have been present or not in state activities. The overall purpose of the research is that by revealing these inadequacies, equal treatment and access for women in these structures might be addressed. (See Kantola, 2007.)

The next area of literature includes feminist IR scholars who begin to move beyond the notion of women as a variable to gender as a broader social construct and analytical category (see also Zalewski, 1998). Rather than focus only on the exclusion or marginalizing of women from the state and the international system, these analyses move toward identifying the gendered nature of the state itself. As Kantola (2007) found, within these studies, the relationship of the state and gender can be reciprocal and constitutive—that the state can be as dependent on gender for its construction and survival as social constructed definitions of gender might be affected by the state. Power is understood as more diffuse and variable, and the state is more than a fixed, abstract thing. It becomes part of a fluid set of processes whereby certain activities involving men and women can work to support or to undermine state authority and position in world politics.

As gendered characteristics of the state are analyzed, a picture of how gender becomes critical to sovereignty emerges. This picture is what informs the third, developing area of literature analyzed by Kantola (2007)—that of the gendered reproduction of the state and debates about state sovereignty and its primacy in the international system. This literature critiqued the traditional understanding of sovereignty, examined debates about whether the state is starting to give way to international and regional organizations, and explored the possibilities that even international and regional organizations are gendered and reproducing state characteristics at a macrolevel.

In general, the traditional understanding of sovereignty has to do with a state's right to govern and its ability to operate independently in the international arena. The concept is understood in political science as related to territories, population, recognition, and authority (Kantola, 2007). Feminist IR has been engaged in deconstructing these aspects of sovereignty for some time, demonstrating that territories and populations are not fixed and indeed the manipulation of gender roles by the state (access to reproductive rights, controls over women's labor, etc.) can play key roles in legitimizing its authority.

However, understanding the complexities of sovereignty is only part of the issue. The intersections of local, national, and global levels are important when talking about states, and the state should be understood in terms of both discursive and structural processes, not just as a unified, unchanging, neutral entity (Kantola, 2007). Gender can matter in the legitimizing of the state and in the gendered and gendering nature of international organizations. Although some of the recent IR literature has suggested that the state is disappearing, other evidence suggests that the state is very much alive and well and in

partnership with international organizations that reinforce gendered characteristics and behaviors. Only through illuminating the continuing gendered and gendering aspects of these structures and institutions can human beings begin to also understand the full complexities of their policies and their effects.

War, Militarism, and Security

In rethinking war, militarism, and security, feminist IR explores to what degree gender as an analytical category can reveal a more complex and complete understanding of violent conflict and its affects on men and women. Sites of inquiry include definitions of war; the impact of war on gender and vice versa; why wars happen and how constructed notions of masculinity and femininity can legitimate militarism and the state; the gendered nature of post-war conditions; and the definitions and policies of security, insecurity, and human security. Several contributions in these areas are noteworthy.

For example, Laura Shepard has recently critiqued the war–peace dichotomies that appear in so much IR literature. She showed how an emphasis on only war or peace can miss the politics of everyday violence, what she described as the violence of in-between times (Shepard, 2009). This violence, which can involve attacks on refugees, recruitment of child soldiers, or the trafficking or forced labor of women, is often left unnamed or unrecognized by the international community.

At the same time, the focus of feminist IR has not just been on women and children as victims of war. As Kelly (2000) and others have noted, in times of war, women can challenge traditional feminine roles, become heads of households, work in nontraditional jobs outside of the home, and even bear arms or serve as soldiers themselves. Women's agency and sense of empowerment can be enhanced by wartime conditions. Furthermore, women's and men's activism in organizing and promoting peace movements can also play out in ways that do not necessarily coincide with stereotypes of masculine and feminine roles and perceptions of war. As El-Bushra (2007) found, rather than viewing conflict as only a violation of women by men, researchers need to think about how men and women are each differently violated by war and what they do about it.

Feminist IR has also recently engaged with scholars who are articulating and analyzing so-called new wars. Mary Kaldor (2006), for example, has pointed to the need to rethink definitions of war and the role of the state. Using Bosnia and Herzegovina as an example, Kaldor examines how the purpose, methods, and financing of war has changed. Where war used to be conducted only by organized states and launched for the purposes of territorial gain or to justify the goals of a particular ideology, many contemporary wars tend to be about identity politics, conducted by decentralized groups, and using methods that

instill fear or create more civilian casualties. Even the financing of such conflicts has changed and become more diffuse and not necessarily dependent on state resources. Peterson (2008) has drawn on Kaldor's work and the literature of international political economy to talk about how these new wars and the economic means of supporting them are gendered as well.

Feminist IR has also explored why wars, new or old, happen and how conflict impacts gender and vice versa. Goldstein's (2001) ambitious work, *War and Gender*, which attempted to bridge traditional IR and feminist IR approaches by using a combination of positivist research and notions of gender as a constructed concept, contributed to the discussion by arguing that connections between war and gender are persistent and interconnected across cultures and time. Killing in war is not natural for either gender, yet the potential for war has been pervasive in all human societies.

Additional feminist IR work has argued for a focus more on militarism and its effects rather than examining just the causes of war. Here, the idea is that militarism itself legitimizes violence as a way to resolve conflict and carries with it the means of redefining gender roles in order to support that behavior (see Kelly, 2000). For example, work has been done on the usage of rape in war, the trafficking of women and children for the purposes of forced prostitution and who benefits, and the impact of posttraumatic stress on both women in the military and on families after troops return home.

The feminist IR literature on security takes research in new directions as well and often interconnects with discussions of conflict and militarism. Again, Tickner (1997, 2002) has been instructive here. She explained that feminist IR scholars have challenged traditional IR's understanding of security as protection of territory or material wealth. Feminist IR sees security more broadly, noting the centrality of the human subject and revolving around protection against all forms of violence, including physical, structural, or ecological. (See also Carpenter, 2005.) To achieve security, one must understand different social hierarchies, the foundational stories and male discourses that support them, and how those same hierarchies influence and may be constructed by various forces in the international system and somehow work to change them.

Influential in some feminist perspectives on security was the publication of the United Nations Development Report (United Nations Development Programme, 1994) that first provided a definition for the term *human security*. The report argues that human security requires the protection of human lives in ways that enhance and support human freedoms and potential. Feminist IR research, although acknowledging the potential of this concept, continues to ask probing questions about whether UN attempts at human security programs fall short and how such policies may still be affected by gender distinctions or limited by Western biases or aspects of the global capitalist economy.

Globalization, Development, and International Political Economy

Feminist IR has also tackled the gendered aspects of globalization, development, and international political economy, redefining these terms and contributing to a wider understanding of men's and women's experiences of political, social, and economic relations. For example, Ruth Paterson (1999) has been an important contributor in this literature. For her, globalization should not be narrowly defined in terms of the international exchange of material goods or the flow of international trade. The study of international political economy should not be limited to the analysis of trade agreements or the work of the International Monetary Fund, the World Bank, or the World Trade Organization or states' political and economic interactions. Globalization, development, and the studies of international political economy need to be conceived in much broader terms and involve the spheres of women's work and activities. Studies should explore to what extent or how gender has mattered and to what degree women in particular may have been disadvantaged or advantaged by certain gendered structures and institutions. Traditional international political economy must be challenged to recognize that the household and private domains where women may dominate are critical in fleshing out connections between economics and politics in all domains (see also Tickner, 2005).

Feminist IR has indeed moved in the direction of focusing on the local experience of women in order to understand larger global economic and political issues. Some interesting and important contributions have included the following: Christine Chin's (1998) book on Filipina and Indonesian female domestic workers in Malaysia, which demonstrated the connections among women's labor, the state, and modernization projects and Elizabeth Prügl's (1999) work, *The Global Construction of Gender: Home Based Work in the Political Economy of the 20th Century*, where she explored numerous links between women's work and economic and political issues.

Transnational Feminism and Human Rights

In the last 5 years, feminist IR has also turned its attention to a critical examination of issues of human rights and the emergence of transnational feminism. Transnational feminism, often used more as a descriptive term than a theoretical framework, refers to the existence of issues, advocacy groups, and organizations across borders and cultures that are interested in the treatment and condition of women (Desai, 2007). The issues that occupy transnational feminism go from a focus on political representation to a concern about economic and social conditions experienced by women in the South and developing or underdeveloped countries as well as in the developed global North. The United Nations has been perceived as the center for

consensus building about these issues, particularly with the creation of key documents like CEDAW (Convention for the Elimination of Discrimination Against Women) and practices such as gender mainstreaming, which advocated for the representation of women in government structures and international organizations for the purposes of drawing women into policy-making processes.

Three edited collections have made significant contributions to the study of these topics: Meyer and Prügl's (1999) *Gender Politics in Global Governance*; *Feminist Politics, Activism, and Vision: Local and Global Challenges*, edited by Ricciutelli, Miles, and McFadden (2004); and *Global Feminism: Transnational Women's Activism, Organizing, and Human Rights*, edited by Ferree and Tripp (2006). A new work by V. Spike Peterson and Anne Sisson Runyon (2010) tackles global gender issues and the importance of understanding gendered lenses on world politics. In these works, feminists investigate what role transnational feminist networks play inside and outside the international system; to what degree international and regional organizations have remain gendered even with the growing consensus about the need for women's participation; and what challenges remain in understanding diversity, multicultural perspectives on human rights, and indeed connections between global and local agendas.

Usefulness of Feminist International Relations and Policy Implications

One of the shared desires among practitioners of feminist IR is that their research be useful to both women and men, whether that be at the level of international organizations or in individuals' day-to-day lives. In the United Nations or the European Union, practical implications can range from the method of collecting data and its interpretation to the usage of those interpretations in determining policy. With gender mainstreaming, for example, intellectual theorizing, advocacy for inclusion, and real-world policy meet. Women have started to be included in some of the bureaucracies, decision-making processes, and programs of these organizations. However, problems still exist in recognizing diversity, acknowledging multicultural and cross-national differences of men and women, and in being self-conscious, reflective, and careful in understanding the gendered nature of policy practices (Carney, 2004).

Along with impacting how data are collected and interpreting policies at either the state or the international level, feminist IR research itself can positively affect individual lives. One interesting example is the work of Maria Mies as described by Tickner (2005). Mies conducted research on sexual violence against women among rural women workers in India. Instead of remaining a distant observer or doing research for only an academic purpose, Mies invited the women of the area who were subjects of the study to participate in research of the researchers as well. All the

results were translated into the local language of Telugu and made available to all. As Mies noted, the process sparked self-confidence among the women and allowed them to work collaboratively together to initiate new solutions and programs to benefit the community as a whole (cited in Tickner, 2005).

Another example of theory and practice coming together is the work being done by McGill University on human security, gender, and peace (Boyd, 2005). Their research and activities have resulted in special refugee programs in Montreal, supporting regional peace initiatives in various parts of Africa, and training for those who work with traumatized victims of war-torn countries worldwide. A central value in their approach is sharing the knowledge gained from groups and translating that knowledge into practical action whenever possible. As feminist IR scholars continue to invite conversation and collaboration with their subjects, no doubt practical implications and creative directions for new research will emerge.

Future Directions

The future directions for feminist IR are numerous, and only a few can be mentioned here. First, although feminist IR theory has created a space for the exploration of a variety of perspectives and topics, more work still needs to be done in examining or even acknowledging non-Western discourses about knowledge, transnational communities, and gender. As Giorgio Shani (2008) suggested, one future consideration for IR might be the exploration of various Islamist or even Sikh viewpoints on universality of culture, community, and politics and how those views intermix with diverse constructions of gender and social behavior.

Another dimension might be further investigation of the emergence of so-called NGO-ization of women's movements and advocacy groups. NGO-ization refers to the proliferation of formal women's NGOs, especially in regions such as the Middle East, sub-Saharan Africa, or even central and eastern Europe. On one hand, these organizations can be viewed as a promising sign of civil society and empowerment of women; on the other, they can be affected by the agendas of international donors, support of UN agencies, and other Western groups, sometimes fragmenting or weakening the possibility of sustained social change (Jad, 2003; see also Kamrani, 2007). Additional analysis might provide an expanded and deeper understanding of the ways these organizations work, where gendered discourses and distinctions may still come into play, and what their roles are in shaping international politics.

Another area of growing research currently involves how queer theory might contribute to feminist IR. Queer theory, as a critical theory of power, sex, and sexuality, can provide an additional lens for examining and understanding social realities and international politics. Here, queer theory and research can point to issues of human rights, variations

of definitions of masculinities and femininities, and how these can enhance understandings of gendered structures and institutions (Kinsella & Holm, 2007).

Finally, continued work in the development of research methods is another area important to both feminist IR and the field of political science as a whole. Expanding a menu of methodological choices with interdisciplinary, qualitative, and quantitative tools can only help to yield more and more sophisticated causal and constitutive understandings of gender and global politics.

Conclusion

This chapter summarized some of the key themes, theories, and methods associated with feminist IR. Like other reviews of feminist IR literature, the purpose here was not to be fully comprehensive or to somehow provide a complete description of how every aspect of feminist IR works. Instead, the intention was to give students of political science a brief sense of the key issues and debates. What students should take away from this overview is that feminist IR has expanded and increased knowledge about gender and global issues. Some of the main contributions include a critique of the methods of research of traditional IR; a reexamination of war, militarism, and security; the state and sovereignty; globalization, development, and international political economy; and finally, continuing investigations of transnational feminism and human rights.

Even though the contributions of feminist scholars have been extensive in the scholarly literature since the 1990s, feminist IR perspectives are still slow to reach undergraduates in the United States. As a recent survey revealed, only about 5% of faculty teaching international relations in the United States devote time in class to discussion of feminist IR, and about 13% of faculty use textbooks on international relations written by women in their classrooms (Jordan, Maliniak, Oakes, Peterson, & Tierney, 2009). Even though academic research and publication on international gender issues has blossomed, classroom discussion of these topics remains limited. It is to be hoped that this review of feminist IR gives students a good place to start their own inquiries about gender and international politics. A list of references and further readings follows.

References and Further Readings

- Ackerly, B., Stern, M., & True, J. (2006). *Feminist methodologies for international relations*. Cambridge, UK: Cambridge University Press.
- Batinic, I. (2001). Feminism, nationalism, and war: The "Yugoslav case" in feminist texts. *Journal of International Women's Studies*, 3(1), 2–23.
- Boyd, R. (2005). Gender and human security issues: Building a programme of action research. *Development in Practice*, 15(1), 115–121.
- Caprioli, M. (2004). Feminist IR theory and quantitative methodology: A critical analysis. *International Studies Review*, 6, 253–269.
- Caprioli, M., & Boyer, M. (2001). Gender, violence, and international crisis. *Journal of Conflict Resolution*, 45, 503–518.
- Carney, G. (2004). *Researching gender mainstreaming: A challenge for feminist IR*. Paper presented at the annual meeting of the International Studies Association, Montreal, QC, Canada.
- Carpenter, C. (2005). Women, children and other vulnerable groups: Gender, strategic frames, and the protection of civilians as a transnational issue. *International Studies Quarterly*, 49, 295–334.
- Chin, C. B. N. (1998). *In service and servitude: Foreign female domestic workers and the Malaysian "modernity" project*. New York: Columbia University Press.
- Desai, M. (2007). The perils and possibilities of transnational feminism. *Women's Studies Quarterly*, 35(3–4), 333–338.
- El Bushra, J. (2007). Feminism, gender, and women's peace activism. *Development and Change*, 38(1), 131–147.
- Enloe, C. (2001). *Bananas, beaches and bases: Making feminist sense of international politics* (Updated ed.). Berkeley: University of California Press.
- Ferree, M., & Tripp, A. (Eds.). (2006). *Transnational women's activism, organizing, and human rights*. New York: New York University Press.
- Fukuyama, F. (1998). Women and the evolution of world politics. *Foreign Affairs*, 77(5), 24–40.
- Goldstein, J. (2001). *War and gender: How gender shapes the war system and vice versa*. New York: Columbia University Press.
- Hooper, C. (2001). *Manly states: Masculinities, international relations and gender politics*. New York: Columbia University Press.
- Jad, I. (2003). *Gender myths and feminist fables: Repositioning gender in the development policy and practice*. Paper presented at the international workshop Feminist Fables and Gender Myths, Institute of Development Studies, Sussex, UK.
- Jonasdottir, A., & Jones, K. (Eds.). (2009). *The political interests of gender revisited: Redoing theory and research with a feminist face*. New York: United Nations University Press.
- Jordan, R., Maliniak, D., Oakes, A., Peterson, S., & Tierney, M. J. (2009). *TRIP survey*. Available at [http://irtheoryandpractice.wm.edu/projects/trip/Final Trip Report 2009.pdf](http://irtheoryandpractice.wm.edu/projects/trip/Final%20Trip%20Report%202009.pdf)
- Kaldor, M. (2006). *New and old wars: Organized violence in a global era*. Cambridge, UK: Polity Press.
- Kamrani, M. (2007). *Keeping it local: Muslim women resisting patriarchy, NGOization, and women's human rights regimes*. Paper presented at the International Studies Association 48th Annual Convention, Chicago.
- Kantola, J. (2007). The gendered reproduction of the state in international relations. *British Journal of Politics and International Relations*, 9, 270–283.
- Kelly, L. (2000). Wars against women: Sexual violence, sexual politics, and militarized states. In S. Jacobs, R. Jacobson, & J. Marchbank (Eds.), *States of conflict: gender, violence, and resistance* (pp. 45–66). London: Zed Books.
- Keohane, R. (1989). International relations theory: Contributions of a feminist standpoint. *Millennium Journal of International Studies*, 18(2), 245–253.

- Kinsella, H., & Holm, P. (2007). *What can queer theory do for you?* Paper presented at the International Studies Association 48th Annual Convention, Chicago.
- Meyer, M., & Prügl, E. (Eds.). (1999). *Gender politics in global governance*. Boston: Rowman & Littlefield.
- Paterson, R. (1999). Towards a feminist international political economy. *Melbourne Journal of Politics*, 26, 1–27.
- Peterson, V. S. (2008). "New wars" and gendered economies. *Feminist Review* 88(1), 7–14.
- Peterson, V. S., & Runyon, A. S. (2010). *Global gender issues in the new millennium*. Boulder, CO: Westview Press.
- Prügl, E. (1999). *The global construction of gender: Home based work in the political economy of the 20th century*. New York: Columbia University Press.
- Prügl, E. (2007). Doing feminist international relations: A guide. *International Studies Review*, 9, 90–92.
- Ricciutelli, I., Miles, A., & McFadden, M. (Eds.). (2004). *Feminist politics, activism, and vision: Local and global challenges*. Toronto, Canada: Inanna Education.
- Shani, G. (2008). Toward a post western IR: The Umma, Khalsa Panth, and critical international relations theory. *International Studies Review*, 10(4), 722–734.
- Shepard, L. (2009). Gender, violence, and global politics: Contemporary debates in feminist security studies. *Political Studies Review*, 7(2), 208–219.
- Sylvester, C. (1994). *Feminist theory and international relations in a postmodern era*. Cambridge, UK: Cambridge University Press.
- Sylvester, C. (2002). *Feminist international relations: An unfinished journey*. Cambridge, UK: Cambridge University Press.
- Tickner, J. A. (1992). *Gender in international relations: Feminist perspectives on achieving global security*. New York: Columbia University Press.
- Tickner, J. A. (1997). You just don't understand: Troubled engagements between feminists and IR theorists. *International Studies Quarterly*, 41(4), 611–632.
- Tickner, J. A. (2001). *Gendering world politics: Issues and approaches in the post cold war era*. New York: Columbia University Press.
- Tickner, J. A. (2002). Feminist perspectives on international relations. In W. Carlsnaes, T. Risse Kappen, B. Simmons, & T. Risse (Eds.), *Handbook of international relations* (pp. 275–291). Thousand Oaks, CA: Sage.
- Tickner, J. A. (2005). What is your research program? Some feminist answers to international relations methodological questions. *International Studies Quarterly*, 49(1), 1–21.
- United Nations Development Programme. (1994). *Human development report 1994: New dimensions of human security*. Available at <http://hdr.undp.org/en/reports/global/hdr1994>
- Waring, M. (1988). *If women counted: A new feminist economics*. San Francisco: Harper & Row.
- Weber, C. (1994). Good girls, little girls, and bad girls: Male paranoia in Keohane, Robert critique of feminist international relations. *Millennium Journal of International Studies*, 23(2), 337–349.
- Wibben, A. (2004). Feminist international relations: Old debates and new directions. *Brown Journal of World Affairs*, 10(2), 97–114.
- Zalewski, M. (1998). Where is woman in international relations? To return as a woman and be heard. *Millennium*, 27(4), 847–867.

42

LEADERSHIP AND DECISION MAKING

TOBIAS VAN ASSCHE

University of Antwerp

Studies of political leadership, which are mostly applied to the foreign policy and crisis management domains, examine how the behavior of individual political leaders can have an impact on policy-making processes and how this can affect what types of decisions are made. This literature draws on psychology to identify personality characteristics that can have an important effect in the political realm and then uses these characteristics as independent variables to explain how they might influence the political process. These characteristics can be very stable, such as personality traits, or can be more volatile over time and content matter, such as cognitions and motives. Individual leaders are one of three forms of decision units that are used to explain types of outcome in foreign policy decision making. Other types of decision units are a single group and a coalition of autonomous actors. This section mainly focuses on political leaders.

After discussing why and when studying political leadership is important, this chapter goes on to discuss three different aspects of a leader's personality that can influence the political process: motives, cognitions, and traits. Next, it describes how some studies have attempted to combine three components to more accurately explain political behavior and outcomes. Finally, areas of future research are identified.

Why Is Studying Political Leadership Important?

In 2003, U.S. President George W. Bush decided to invade Iraq and forcefully remove its authoritarian president, Saddam Hussein, from power. This policy was very different from the first Gulf War since the United States accomplished its objective with the help of a much smaller coalition of countries that did not involve some of the traditional allies, such as France and Germany, and without support of the United Nations Security Council. Despite these international constraints and weak evidence concerning the *casus belli* (cause of war)—weapons of mass destruction—the president and his advisors strongly believed that invading Iraq was necessary and feasible. What the policymakers did not anticipate was that it would draw the United States into a prolonged conflict in Iraq, that weapons of mass destruction would not be found, and that it would drain resources from the war in Afghanistan, which was being conducted simultaneously. What makes this case interesting for people studying political leadership is that it was a war of choice, not one of necessity—a choice made by the president. In a case such as this, one therefore cannot help but wonder if a different leader would have taken an alternative course of action. What

would have happened if the Supreme Court had declared Al Gore the winner of the 2000 presidential elections? Would he also have chosen to invade Iraq, or would he have selected a different course of action to deal with Saddam Hussein?

The 2003 Iraq War is just one of a large number of international cases where the person in charge appears to have played a vital role in its outcome. Robert Kennedy, the U.S. Attorney General and brother of President John F. Kennedy, for example, stated about the EXCOM—the group of high-level policymakers who dealt with the Cuban missile crisis—that “if six of them had been President of the U.S., I think the world would have been blown up” (Steel, 1969, p. 22). One could also argue that Saddam Hussein played an important independent role in his country’s decision to invade Kuwait in 1990 (Post, 2003b), or that Woodrow Wilson’s self-defeatist behavior influenced the outcome of the Versailles Peace Conference, which brought World War I to a close (George & George, 1956).

Although political leaders make important decisions on a daily basis and are directly linked to decisions that the administration makes, it is also important not to fall into the trap of attributing every international event merely to the personality of the participating countries’ elites. There are many patterns and specific situations in international relations when one does not need to look at who was in charge to explain the outcome. The personalities of the U.S. presidents during the cold war did not affect the stalemate between NATO and the Warsaw Pact, and it is not necessary to know who is in charge in Britain and France to explain why these countries no longer fight each other. The personality of the political leaders does not matter in these conditions because their beliefs and perceptions can largely be explained by the situation. Most leaders would act similarly in these conditions.

In studying political leadership, we are particularly interested in examining the various aspects of the decision-making process, relying on psychological frameworks within an institutional setting, in cases where it not only provides us with a more in-depth understanding of what happened but also makes it possible to explain significant additional variation in outcomes. We want to examine cases where the leader really made a difference.

The need to analyze political leadership has increased considerably following the end of the cold war, since there no longer is an agreement on the way in which the world is structured. Because of the threat of World War III and mutually assured destruction during the cold war, it was clear that no leader would significantly challenge the status quo, so conflicts were limited to the peripheral areas such as Vietnam and Korea. Since the fall of the iron curtain, and the Soviet Union in particular, this is no longer the case. There is now much more room for interpretation, innovation, misunderstanding, and miscommunication than when a bipolar world system dominated

international interactions (Hermann & Hagan, 1998). For instance, there is no agreement on how big the threat of Islamic terrorism is; it is unclear to what extent China, Russia, and the European Union are threatening the United States in its hegemonic role; and countries such as Iran and North Korea are very unpredictable in their pursuit of nuclear weapons. How the elites interpret this unstable environment is thus increasingly important.

Understanding the influence of political leaders is not something that is only interesting to academics. Branches of the U.S. government such as the CIA often draw on analyses of elites to assist in the policy-making process (Post, 2003a). One classic example is how U.S. President Jimmy Carter relied on personality profiles of Israeli Prime Minister Menachim Begin and Egyptian President Anwar Sadat during the Camp David negotiations, which led to a peace treaty between the two countries. The profiles, constructed by leading political psychologists, warned the president that the personalities of the two leaders conflicted significantly; Begin was very detail oriented, while Sadat focused more on the big picture, which could significantly complicate negotiations. Based on this information, Carter designed a strategy that eliminated direct contact between the two protagonists and allowed him to act as a middleman. To this day, experts in political leadership cooperate with other academics such as cultural anthropologists to advise the U.S. Government.

When Is the Decision-Making Process Important?

Since it is not always equally necessary to focus more narrowly on the decision-making process to understand why a country acts the way it does in the international arena, it is important to identify under which conditions leadership and the decision-making process are most likely to have an autonomous impact. As Post (2003a) states, the goal is to distinguish between situations that conform to the “covering-law generalization from a structural theory about a universe of cases or deviates from it due to the operation of intervening causal mechanisms between structural conditions and decision outcomes” (p. 64).

Scholars have been able to identify certain conditions in which the decision-making process is more likely to have an important independent impact. Hagan (2002) states that the decision making process is particularly important when the leaders (a) face real uncertainty in responding to international threats; (b) are confronted with trade-offs across competing goals, including that of retaining power; and (c) operate in decision structures in which political authority is quite dispersed and fragmented. Examining decision making in these conditions is important because a lack of information and goal certainty makes it very difficult to determine the rationally optimal

course of action. The outcome will thus strongly rely on “how leaders perceive and interpret the threats based on their belief systems” (Hagan, 2002, p. 11). One could thus expect that in these cases different types of leaders might react differently.

One specific situation in which the decision-making process can play an important role is in a crisis (Hermann, 1976). A crisis is defined as a situation in which the policymakers perceive uncertainty, a threat to core values, and time pressure. In these conditions, policymakers are forced to make quick important decisions, often with limited information. Because they are not necessarily able to have access to all information and have to deal with issues such as stress (Janis, 1982), psychological factors and the nature of the decision-making structure can become important (Vertzberger, 1990). This was exemplified in Jervis’s (1976) groundbreaking work in which he discusses how perception and misperception can influence international politics.

It is important to stress that political leadership can be a necessary factor to explain why a certain event happened, but it is never sufficient. A political leader can be important because of the way in which this person interprets the environment, but it is the context and conditions that matter primarily. Elites only determine to what extent and in which fashion cues from the environment are interpreted.

Political Leaders as a Decision Unit

The personality of political leaders is not the only factor that can influence how decisions are made. The literature on foreign policy decision making, building on classics such as Graham Allison’s (1971) *The Essence of Decision*, distinguishes among three types of decision units: a single group, a coalition of autonomous actors, and a powerful leader (Hermann, 2002). It is important to distinguish among these three types, because different factors influence how these distinct decision units come to determine a policy. The first type is a single group. This is a group composed of two or more policymakers in which all members are necessary to make the decision to commit resources. Examples of a single group are the EXCOM during the Cuban missile crisis and President Lyndon Johnson’s inner cabinet during the Vietnam War. Since the focus is on how a group of people comes to a decision, the main variables used to explain how this decision unit selects a policy are derived from the study of group behavior in social psychology. This is because in a single group, decisions are made through a collective, interactive, decision process, in which all members participate. The focus is therefore on factors such as group pressures and the presence of a minority.

A second decision unit is a coalition of autonomous actors. This decision unit is composed of a group of independent actors who need to work together to come to a decision. Although the group members can share common

goals, their primary loyalty is not toward the coalition. They have their own constituency that they are accountable to. In the United Nations Security Council, for example, the member countries do not act as a single group. Their primary loyalty does not lie with the United Nations but their respective countries’ self-interests. The same can be said for the political parties that compose a coalition government. This decision-making unit focuses more on the types of decision rules and factors that can help actors with different positions come to an agreement.

The focus in this chapter is on political leadership, or situations where an individual is obliged to, or chooses to, take the authoritative decision. This type of decision unit is predominant in countries or international organizations where one individual is vested with the authority to commit or withhold the resources of the government with regard to foreign policy matters (Hermann, 2002). Here, the focus is on individual leaders and how their personalities can influence the decision-making process and policy outcome. Who is in power can affect what information is taken into account, what factors motivate a decision, and how the advisory system is structured (who takes part in which format; Hermann & Preston, 1994), and so on. The predominant leaders are particularly likely to be found in authoritarian countries and dictatorships such as Mugabe in Zimbabwe and Lukashenko in Belarus but can also be important in countries such as the United States, where the president enjoys a lot of independent power in the foreign policy realm, among others. The systematic influence of leadership can also be detected in other cases such as international organizations and conditions where the leader is first among equals, but in these conditions, the individual impact will be a lot less direct and therefore more difficult to detect or measure. They can, for example, attempt to set the agenda or try to persuade other important actors, but their policy preferences do not deliver direct results.

Psychoanalytic Studies of Political Leadership

The study of political leadership emerged in the early 20th century with the rise of psychoanalysis, most famously associated with the work of Sigmund Freud, in which an important figure’s behavior is explained by investigating his or her formative period. Scholars started using these Freudian techniques to analyze why an individual leader acted the way he or she did at a certain time, or to explain his or her overall behavior, in what came to be called psychobiographies. The root causes in these studies are often the individual’s relationship with parents, childhood traumas, or formative events. A classic example is Juliette George and Alexander George’s (1956) study of Woodrow Wilson.

There have also been attempts to move beyond explaining individual behaviors using psychoanalysis and move toward a more comparative approach. An example of this is James Barber’s (1992) study of the U.S. presidential

character, which was originally published in 1972. He claims that the U.S. presidents can all be categorized as being active or passive in the amount of energy they put into the job and can be either positive or negative based on the satisfaction they get from the position. A combination of these two factors then determines success or failure in office. Barber argues, for example, that an active–negative president such as Hoover, Johnson, and Nixon is the most dangerous because these types of leaders tend to be compulsive, while an active–positive president such as FDR, Clinton, and Carter is most likely to succeed because these types of leaders tend to be more adaptive.

Although Barber’s model remains extremely popular in classrooms, psychoanalysis is no longer widely used to explain political behavior. This is because this method is not focused on creating testable and generalizable hypotheses but instead wants to explain one individual’s behavior in very specific circumstances. Even if more psychoanalysts wanted to follow Barber and create broad, generalizable theories, the reliance on creating a narrative out of historical material is also not conducive to explaining political behavior and outcomes. As Runyan’s (1981) study of why the Dutch painter Van Gogh might have cut off his ear demonstrates, psychoanalytical theories are also impossible to falsify, since there can be many possible explanations why someone behaved the way he or she did. Scholars also often do not have direct access to the leaders, which means they have to rely on documents and secondary sources, which can be a problem when trying to create reliable measures (Houghton, 2009).

Personality and Politics

Political psychologists followed the evolution in psychology and no longer focus solely on psychoanalysis but instead investigate different specific components of personality to analyze political behavior. Winter (2003c) defines personality as “an array of capacities or dispositions that may be engaged, primed, or brought forward depending on the demands of the situation and a person’s own ‘executive apparatus’” (p. 12). Winter compares personality to a personal computer with some relatively fixed hardware characteristics and software applications that can be opened or closed by the operator and affect the arousal and weighting of leaders’ goals and preferences, as well as conflicts and confusion among different goals.

Someone’s personality is important because it acts like a filter through which information is processed and interpreted: It affects how someone interprets information from the environment, how he or she will react to the stimuli, his or her persistence and endurance, or how someone manages factors such as stress and emotions.

As the discussion on psychohistory demonstrated, one of the stumbling blocks to studying political leadership has always been that the investigators do not have access to the individuals whom they want to study. The reason for this

is that presidents, prime ministers, and other important foreign policy actors, such as Barack Obama and Vladimir Putin, often do not have time, or are not interested, to complete the necessary psychological tests or clinical interviews, and even if they had the time, they might not want to participate since the results might be politically damaging if made public. Consider, for example, the implications if through direct tests, psychologists discover that a prime minister is clinically depressed. People might start questioning his or her ability to lead, which would likely be career ending.

To resolve the accessibility problem, scholars developed a number of techniques to derive the personality components at a distance. These methods, which are based on existing psychological measures and tests, rely on content analysis of text material, such as speeches and interviews, because word use can not only demonstrate conscious attempts to relay a message but also encompass unconscious factors such as beliefs, motives, and cognitive structures. The validity and reliability of these techniques has been demonstrated through systematic research by, for example, comparing the results to those of expert panels (Winter, 2003a).

Many of these techniques can now be coded automatically, using computer software. For example, Margaret G. Hermann (1980) developed at-a-distance measures of several beliefs and interpersonal styles, such as nationalism, authoritarianism, and self-confidence (which will be discussed further in a following section). Michael Young subsequently created ProfilerPlus, a program that codes the text automatically using word libraries. The source material—the type of text—that is used for each method can differ significantly. In some cases such as leadership trait analysis, the scholars are looking for spontaneous text, like interview responses, while in other cases, they can rely on more scripted material such as speeches. Similar techniques have also been developed to code other frameworks such as the operational code, which will also be discussed later.

Single Personality Variables

The following section focuses on the three different components of personality: (1) motives, (2) cognitions and beliefs, and (3) temperament and interpersonal traits, which are also often labeled *personality traits* (Shafer, 2000). (Winter, 2003c, adds the situational context as a fourth component, but this is less generally accepted.) Each part briefly defines the component, addresses how this aspect of someone’s personality can affect the way a leader behaves, discusses some important measures that are used, and mentions a number of specific findings.

Motives

One group of factors that can influence how someone will behave in the political realm is his or her motives.

These are the different classes of goals toward which people direct their behavior. As Winter (2003b) states, motives influence how leaders construe the leadership role. Motives influence perceptions of opportunity and danger, they affect the accessibility of different styles and skills, and they determine sources of leadership satisfaction, stress, frustration, and vulnerability. For example, two students with equal intelligence and skills can make different decisions based on what drives them. If one of them is motivated because he or she wants to be successful in his or her studies while the other is more interested in being popular among peers, their attitudes toward school are likely to differ significantly. The interpersonally motivated student is more likely to be swayed to go to a party the day before a test while a task-driven student is more likely to stay home to study. As a result, their exam scores will likely differ, despite the two students' equal intelligence.

Motives are not constant; they can change because of external incentives and internal dynamics and can be subject to distortion, deception, and rationalization. Goals can be accomplished, people can be disappointed, and new goals can emerge. Although there are many different motives that can drive people to act in a certain way, studies in political leadership primarily focus on three: (1) need for power, (2) achievement, and (3) affiliating themselves with others. As with many other personality variables, motives are often measured at a distance through content analysis of verbal behavior.

A first important motive is termed *need for power*. Someone who scores high on this variable is concerned about having an impact, control, or influence over another person, group, or the world at large. A high score on need for power does not mean that this individual is necessarily a Hitler, Napoleon, or comparable to any villain in a James Bond movie. Instead, these people can have qualities that can be beneficial as well as negative for their leadership function. According to Winter (2003b), people who score high on need for power tend to be more adept at building alliances with others, and they actively define the situation themselves, encourage people to participate, and try to influence others. During negotiations, they tend to be exploitative and aggressive and tend to rely on political experts for advice, instead of friends. Hermann (1980) found that among 45 world leaders, power-motivated leaders tended to be more independent and confrontational. Winter (1980) also states that power motivation is associated with involving the country in war.

Another motive that can impact political behavior is termed *need for achievement*. This motive refers to a person's concern about achieving excellence (Winter, 2003b). People who score high on this variable tend to be rational calculators, pursuing their self-interest. Although achievement-motivated individuals do well in business, they do not necessarily function well in the political domain. This is because in dealing with a problem, once they have established the best solution, they tend to want to

push this plan through, even if others disagree. In dealing with a financial crisis, for example, some believe that large-scale government intervention is necessary while others believe the market will balance itself. In such conditions, an achievement-oriented leader can alienate part of his or her constituency by pushing forward a divisive program.

The motive termed *need for affiliation* measures to what extent someone is concerned about establishing, maintaining, or restoring friendly relations among persons and groups (Winter, 1991). Leaders who score high on this motive are more focused toward others. Their circle is often small and consists of people whom these leaders agree with and whom they like. These types of leaders tend to be more peaceable and cooperative (Hermann, 1980)—so long as they are surrounded by like-minded others and do not feel threatened. They are also more vulnerable to the influence of self-seeking subordinates and scandal. Nixon, for example, scored high on affiliation as well as achievement (Winter & Carlson, 1988).

Cognitions and Beliefs

A second way in which a leader's personality can influence the decision-making process is through his or her cognitions, beliefs, or both. This includes a wide variety of mental representations, schemas, models, categories, beliefs, values, and attitudes (Winter, 2003a). This category analyzes ways in which individuals view various friendly and hostile groups, different social systems, and themselves. It also encompasses how these leaders interpret, structure, and retrieve information, as well as leaders' general beliefs about the nature of politics and the world in general. Cognitions and beliefs can play a very important role because they determine how information is processed and which information is selected. For example, someone who is convinced that the world is an evil place where everybody is out to get him or her is more likely to accept information that supports this worldview than someone who believes the international realm is more peaceful and cooperative. Beliefs tend to be fairly fixed but can be affected by persuasion. An individual who grew up fearing the Soviet Union because of the cold war is very unlikely to quickly change his or her opinion about Russia and its people. He or she can, however, change beliefs by, for example, moving to Russia or having to work with a lot of Russians on a daily basis.

One of the most studied cognitive features is conceptual complexity. Do leaders process information in simplistic ways, or do they recognize different points of view? Do they view the world in black and white alternatives, or are they able to recognize many different shades of gray? Conceptual complexity can be treated as a stable personality trait. This characteristic can lead to leadership success, affecting, for example, the length of tenure in high office and can reduce stress during crises (Wallace & Suedfeld, 1988). Bill Clinton, for example, scores fairly

high on conceptual complexity compared to other world leaders (Hermann, 2003b).

Studies have also examined how the context might influence complexity or how people with different levels of complexity react to different conditions (Suedfeld, Guttieri, & Tetlock, 2003). Suedfeld, Tetlock, and Streufert (1992) refer to this interaction as integrative complexity. High scores on integrative complexity were found to be related to peaceful resolution (versus escalation) of international conflict (Tetlock, 1977). It does not mean, however, that high complexity leads to better decisions. Chamberlain, who was comparatively complex during the Munich Conference, was outmaneuvered by Hitler in spite of the latter's low level of complexity (Suedfeld et al., 2003).

An important tool for measuring a leader's beliefs is the operational code (George, 1969). This is a set of philosophical beliefs about the nature of political life (harmony or conflict), the predictability and controllability of political outcomes, and instrumental beliefs about the best way to pursue goals and calculate risks. These beliefs function as guides to political decision making, especially under conditions of relative uncertainty about the consequences of choosing one course of action rather than another (Walker, 1990). These beliefs vary in degree over time and for different issue areas (Walker, Shafer, & Young, 2003). Renshon (2008), for example, demonstrated with his analysis of President George W. Bush that role changes and traumatic events caused major shifts in his beliefs. This technique is particularly useful for gaining an understanding of new actors about whom we know little, as was the case when Vladimir Putin replaced Boris Yeltsin as Russian president. Walker, Schafer, and Young (1998) have developed objective quantitative methods for assessing operational codes based on the verbs in context approach.

A number of other beliefs and interpersonal style variables have also been studied that can affect how leaders act in the political realm. Hermann (1980), for example, developed at-a-distance methods to measure level of nationalism, distrust of others, belief that one can control events, and self-confidence. Hermann found that distrust and nationalism were linked to expressions of strong negative affect toward other countries and with low levels of resource commitment in foreign relations. She also found that leaders with high levels of self-confidence are more immune to incoming information and feel less compelled to adapt to the nature of the situation.

Traits

A third aspect of a person's personality is a leader's traits. Although many psychologists use the term *trait* to refer to all personality variables, others define this term as the public, observable elements of personality. Without necessarily knowing someone's beliefs, motivations, or intellectual capacities, it is possible to make some statements about an individual, purely based on what one sees. In the popular TV series *Friends*, for example, the

character Rachel Green, played by Jennifer Aniston, can be described as being self-absorbed, outgoing, and image conscious, while Monica Geller, played by Courteney Cox Arquette, is obsessive, competitive, and neurotic. Scholars mostly agree that there are five big traits that are considered fundamental and are therefore most frequently used by political psychologists. They are (1) extraversion or surgency, (2) agreeableness, (3) conscientiousness, (4) emotional stability or neuroticism, and (5) openness to experiences (Winter, 2003a). Although most concur that these are the five most salient traits, there is considerable discussion as to the content and structure of each factor.

These variables are often coded by asking historians, or other individuals that have studied the leader, to fill in questionnaires that probe for the various traits, using, for example, adjective check lists or by content analyzing descriptions of leaders by other policymakers. Traits resembling extraversion and openness have been linked to a variety of measures of presidential performance (Simenton, 1988), and other studies found a correlation between openness and greatness (Rubenzer, Faschingbauer, & Ones, 2000).

Multivariate Approaches

The previous section treated the three major categories of personality—motives, cognitions, and traits—separately, focusing on how these different components can impact a leader's decision. Very often, however, behavior is not determined by one sole factor but is instead an interaction between different parts of someone's personality. Motives might drive you to do something, but how you interpret the situation determines whether you even observe that there is a problem or not. Some studies have attempted to combine different personality aspects to examine political behavior.

One attempt to combine the various personality dimensions is Hermann's (1987) Leadership Trait Analysis (LTA). This approach combines seven personality traits to create an overall leadership profile. These are conceptual complexity, belief that the leader can control events, self-confidence, need for power, task focus, distrust of others, and in-group bias. Combinations of these variables are then used to answer three questions: Is the leader open or closed to information, is he or she task or relationship focused, and will the individual respect or challenge constraints? How the leader scores on these questions then determines which of eight leadership styles a leader has: expansionistic, evangelistic, actively independent, directive, incremental, influential, opportunistic, or collegial. Bill Clinton, for example, is considered to be a collegial leader: He generally respects constraints, is open to information, and is motivated by both solving the problem and keeping morale high (Hermann, 2003b).

Hermann and Preston (1994) have also linked leadership style to the way in which U.S. presidents organize their advisory systems: whether it is formalistic, competitive,

or collegial. Kille (2006) also combined different LTA variables to determine how secretaries general of the United Nations will approach their positions. He demonstrates that whether the leader is a manager, a strategist, or a visionary will determine how the individual will use his or her agenda-setting power, resolve disputes, and approach possible UN interventions.

Future Directions

So far, studies in political leadership have been very successful in determining that leadership matters (Hermann & Hagan, 1998). Research has found a number of aspects of an individual's personality that can determine politically important behaviors such as the selection of advisors, the processing of advice, risk-taking propensity, and negotiating styles. Research has also found a relationship between some personality aspects and broad foreign policy outcomes such as going to war. Still, leadership as an independent variable in international relations will always be limited in what it can explain since it will always to some extent be influenced by the environment. A leader can be very war prone and can push his or her country toward attacking another nation. Nevertheless, this does not mean that violence will necessarily occur, since other countries might intervene, or domestic factors such as other institutions or popular outcries might still force the leader to change his or her mind. Leadership can thus never by itself fully predict exact situations (Winter, 2003c).

The study of political leadership is a constantly evolving field. Next to working on making the current variables more reliable and valid, scholars are also examining how the different components that have been established might relate to each other and how they might interact in systematic ways. Researchers are also constantly expanding into new areas. They are investigating new personality components such as creativity and impulsivity. The literature is also gradually expanding into examining how emotions such as anxiety and anger influence decision making (Post, 2003a).

Next to finding new and improved measures of personality components, the literature also needs to examine new areas in which leadership can have a significant independent impact. So far, most of the studies have focused on the most likely cases of leadership or examined individuals where the impact on the political realm would be most obvious, such as presidents, authoritarian leaders, or both, or the studies have focused on situations where leaders were forced to take autonomous decisions such as during international crises. One area in which political leadership has yet to expand significantly is situations where the leadership effect is not as direct. Kille and Scully (2003) and Kille (2006), for example, study the secretaries general of the United Nations and the presidents of the European Commission and look at how their leadership styles can be important, despite not having any true formal power that

comes with their positions. Political leadership can also examine cases where leaders such as presidents and prime ministers can play a role outside the area where they have traditionally held most power. One example is the U.S. president's State of the Union address. Do presidents' motivations and openness to information influence whether these individuals will actively try to set the agenda or be more reactive to pressures from the media and public opinion?

The rapidly advancing and broadly available technology also provides great opportunities to advance the knowledge of political leadership in the near future. Websites such as LexisNexis and the World News Connection, among others, allow scholars easy access to a wide variety of source material such as speeches and interviews. This material is often translated into English and can be accessed within days, hours, or in some cases even minutes of the event. Computer-assisted content analysis techniques also allow scholars to code this material in significantly less time, without endangering validity or reliability. This increased availability of data, as well as the ability to code much faster, present the opportunity to do comparative studies of leadership on a much larger scale. These studies can either rely on a much larger sample of political leaders or look at how the personality factors of an individual evolve across time or in different conditions. For example, it is now feasible for scholars with a limited budget to create leadership profiles of all members of the U.S. Congress or other national or international bodies.

The importance of political leadership is likely to continue in the future, since the instability in the world shows no sign of changing and national and international crises can erupt at any time. Scholars will thus continue to have a plethora of cases and individuals to examine and better material and techniques to accomplish this. These results will then, it is to be hoped, be used not only to advance science but also to help design strategies that will avert disasters and help resolve conflicts on a global level.

References and Further Readings

- Allison, G. T. (1971). *The essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.
- Barber, J. D. (1992). *The presidential character: Predicting performance in the White House* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- George, A. L. (1969). The "operational code": A neglected approach to the study of political leaders and decision making. *International Studies Quarterly*, 13(2), 190-220.
- George, A. L., & George, J. L. (1956). *Woodrow Wilson and colonel house: A personality study*. New York: John Day.
- Hagan, J. D. (2002). Does decision making matter? Systemic assumptions versus historical reality. In M. G. Hermann (Ed.), *Leaders, groups, and coalitions: Understanding the people and processes in foreign policymaking* (pp. 5-46). London: Blackwell.

- Hermann, M. G. (1976). Circumstances under which leader personality will affect foreign policy: Some propositions. In J. Rosenau (Ed.), *In search of global patterns* (pp. 326-332). New York: Free Press.
- Hermann, M. G. (1980). Explaining foreign policy behavior using the personal characteristics of political leaders. *International Studies Quarterly*, 24(1), 7-46.
- Hermann, M. G. (1987). *Handbook for assessing personal characteristics and foreign policy orientations of political leaders*. Columbus: Ohio State University, Mershon Center for International Security Studies.
- Hermann, M. G. (2002). How decision units shape foreign policy: A theoretical framework. In J. D. Hagan & M. G. Hermann (Eds.), *Leaders, groups, and coalitions: Understanding the people and processes in foreign policy decision making* (pp. 47-82). Malden, MA: Blackwell.
- Hermann, M. G. (2003a). Assessing leadership style: Trait analysis. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 178-214). Ann Arbor: University of Michigan Press.
- Hermann, M. G. (2003b). William Jefferson Clinton's leadership style. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 313-323). Ann Arbor: University of Michigan Press.
- Hermann, M. G., & Hagan, J. D. (1998). International decision making: Leadership matters. *Foreign Policy*, 110, 124-37.
- Hermann, M. G., & Preston, T. (1994). Presidents, advisers, and foreign policy: The effects of leadership style on executive arrangements. *Political Psychology*, 15(1), 75-96.
- Houghton, D. P. (2009). *Political psychology: Situations, individuals, and cases*. New York: Routledge.
- Janis, I. L. (1982). *Victims of groupthink*. Boston: Houghton Mifflin.
- Jervis, R. (1976). *Perceptions and misperceptions in international politics*. Princeton, NJ: Princeton University Press.
- Kille, K., & Scully, R. M. (2003). Executive heads and the role of intergovernmental organizations: Expansionist leadership in the United Nations and the European Union. *Political Psychology*, 24(1), 175-198.
- Kille, K. J. (2006). *From manager to visionary: The secretary general of the United Nations*. New York: Macmillan.
- Post, J. M. (2003a). *The psychological assessment of political leaders*. Ann Arbor: University of Michigan Press.
- Post, J. M. (2003b). Saddam Hussein of Iraq: A political psychological profile. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 335-366). Ann Arbor: University of Michigan Press.
- ProfilerPlus: <http://socialscience.net/Default.aspx>
- Renshon, J. (2008). Stability and change in belief systems: The operational code of George W. Bush. *Journal of Conflict Resolution*, 52(6), 820-849.
- Rubenzon, S. J., Faschingbauer, T. R., & Ones, D. S. (2000). Assessing the U.S. presidents using the revised NEO personality inventory. *Assessment*, 7(4), 403-420.
- Runyan, W. M. (1981). Why did Van Gogh cut off his ear? The problems of alternative explanations in psychobiography. *Journal of Personality and Social Psychology* 40(6), 1070-1077.
- Schafer, M. (2000). Issues in assessing psychological characteristics at a distance: An introduction to the symposium. *Political Psychology*, 21(3), 511-527.
- Simenton, D. (1988). Presidential style: Personality, biographical and performance. *Journal of Personality and Social Psychology*, 55, 928-936.
- Steel R. (1969). Endgame. *New York Review of Books*, 12, 15-22.
- Suedfeld, P., Guttieri, K., & Tetlock, P. E. (2003). Assessing integrative complexity at a distance: Archival analyses of thinking and decision making. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 246-270). Ann Arbor: University of Michigan Press.
- Suedfeld, P., Tetlock, P. E., & Streufert, S. (1992). Conceptual/integrative complexity. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 393-400). Cambridge, UK: Cambridge University Press.
- Tetlock, P. E. (1977). Identifying victims of groupthink from public statements of decision makers. *Journal of Personality and Social Psychology*, 37(8), 1314-1324.
- Vertzberger, Y. I. (1990). *The world in their minds: Information processing, cognition, and perception in foreign policy decision making*. Stanford, CA: Stanford University Press.
- Walker, S. G. (1990). The evolution of operational code analysis. *Political Psychology*, 11(2), 403-418.
- Walker, S., Shafer, M., & Young, M. (1998). Systematic procedures for operational code analysis: Measuring and modeling Jimmy Carter's operational code. *International Studies Quarterly*, 42(1), 175-190.
- Walker, S., Shafer, M., & Young, M. (2003). Profiling the operational codes of political leaders. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 215-246). Ann Arbor: University of Michigan Press.
- Wallace, M. D., & Suedfeld, P. (1988). Leadership performance in crisis: The longevity complexity link. *International Studies Quarterly*, 32(4), 439-451.
- Winter, D. G. (1980). Measuring the motives of southern African political leaders at a distance. *Political Psychology*, 2(2), 75-85.
- Winter, D. G. (1991). Measuring personality at a distance: Development of an integrated system for scoring motives in a running text. In A. J. Stewart, J. M. Healy Jr., & D. Ozer (Eds.), *Perspectives in personality: Approaches to understanding lives* (pp. 59-89). London: Jessica Kingsley.
- Winter, D. G. (2003a). Assessing leaders' personalities: A historical survey of academic research studies. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 11-38). Ann Arbor: University of Michigan Press.
- Winter, D. G. (2003b). Measuring the motives of political actors at a distance. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 153-177). Ann Arbor: University of Michigan Press.
- Winter, D. G. (2003c). Personality and political behavior. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *Oxford handbook of political psychology* (pp. 110-145). New York: Oxford University Press.
- Winter, D. G., & Carlson, L. (1988). Using motive scores in the psychobiographical study of an individual: The case of Richard Nixon. *Journal of Personality*, 56(1), 57-103.

43

BALANCE OF POWER

BRIAN R. KING

Muskingum University

Balance of power is a concept within the realm of international relations that stretches back centuries in both theory and practice and is still among the prevalent topics of debate within contemporary political science. These centuries of historical perspective and scholarship, however, have served only to intensify the debate over the merits of balance of power theory.

There are many ways in which the term *balance of power* has been used in theory or in practice (Claude, 1962; Sheehan, 1996), and this variety of approaches to the concept demonstrates that the term is often used so freely as to potentially confuse rather than clarify its meaning (Sheehan). Despite this diversity, however, nearly all of these definitions center on the same general principles and assumptions and boil down to the central assertion that nation-states will ally with one another in order to create an equality of capabilities between opposing alliances that serves to preserve peace at the international level. Some scholars have attempted to codify the formal assumptions, conditions, and criteria for labeling an arrangement as a balance of power system, and perhaps the best known of these is Kaplan (1957). Based on theoretical modeling, he delineated six assumptions that had to be accepted, then outlined his six fundamental rules for a balance of power system. Although these and other attempts to formalize the process are aspiring to help the field of study and policymakers alike, the problem is that

balance of power systems in practice neither cohere to all of the assumptions nor follow all of the rules set out by any given treatise on the subject. The global environment and the myriad of other variables are not static but rather are fluid in nature and are therefore difficult to prescribe (Claude, 1962).

Some scholarship has focused on the dynamic that exists between the main balance of power system and various local balances that exist as subsets of the main system, including the roles played by dominant and submissive balances within the overall system (Bull, 1977). Though there is no consensus concerning the full extent of the relationship between these smaller balances and the overall arrangement, it is generally agreed that these play an important role in understanding the true dynamic of any balance of power arrangement.

Core Concepts and Dynamics of the Process

Balance of power theory asserts that nation-states, as rational actors in an anarchical international system, will accumulate power in pursuit of national survival and will concurrently strive to make certain that no other actor accumulates sufficient power to threaten their own security. These assertions are firmly rooted in realist assumptions.

Morgenthau (1948, 1978), the father of 20th-century classical realism, contends (among other things) that nation-states are the primary actors on the world stage, that nation-states are rational actors who make rational decisions, and that states relentlessly pursue power and capabilities in order to protect and promote their national interest (i.e., survival). States in the international system possess varying levels of power, and as such, strong states are perceived to be a threat to the security of weaker states. These weak states may thus find it necessary to enlarge their own military capabilities, known as internal balancing. When this is not an option, whether for economic or other reasons, weak states may ally with other actors in the system, known as external balancing, so that their combined power can sufficiently match that of an adversary or an alliance of adversaries, thus creating a balance among competing powers on the international stage that makes aggression much less likely. The security regime that results is an international system of distrust and competition among nation-states, all of whom could potentially be ally or adversary toward one another depending on a state's security needs and the current balance among states and alliances. This dynamic is nothing new but rather has been a constant and recurring reality in international relations so long as there have been independent actors in an anarchical international system.

Balances occur as an outgrowth of the anarchical nature of the system, a fundamental assumption of the realist paradigm. Nation-states, the principal actors in this system, are ultimately unbound by any rulers or regimes above them, since any such organization or agreement would lack sufficient sovereign authority to compel specific state behavior (Waltz, 1979). The primary goal of each state is survival, which is pursued through the accumulation of power. From the realist perspective, power is generally gained by virtue of military capability, while economic capability is important primarily for its ability to gain additional military power. At this point, two additional realist assumptions come into play. First, the pursuit of power can be seen as something of a zero-sum game at the international level, meaning that the acquisition of power by one state comes at the cost of others in the system. In other words, power being a relative concept, gains by another actor will be perceived by potential adversaries as a relative loss for themselves. Second, aggression against weaker states is assumed to be a likely behavior of stronger states pursuing greater power and capability. In fact, Morgenthau (1948) contends that individual states seek superiority, not balance, and that the concurrent competition for superiority between states is what brings the balance itself.

Both of these assumptions being true, it stands to reason that the existence of weakness—and thus imbalance—in the system will lead to aggression on the part of the strong state, thus threatening the security of all actors. To prevent this, weaker states will put aside their own differences and

form an alliance to protect their own security and prevent the stronger state from becoming hegemonic. It is the perceived threat of hegemony that matters. In this light, even traditional allies and states that have no history of conflict will engage in balancing behavior in order to prevent the ascension of a hegemon in the system (Waltz, 2000). Put another way, hegemony actually compels the process by which balance results (Waltz, 1993).

The realist paradigm, dominant in its influence among scholars and policymakers alike in the early days of the cold war, had by the 1960s and 1970s become the target of increasing criticism from those who saw the approach as both theoretically and methodologically flawed. Out of this criticism came Waltz's (1979) *Theory of International Politics*, which redefined the paradigm as structural realism, or neorealism. According to Waltz, the balance of power concept was still the central organizing principle of the anarchical international system. States, in order to survive, are left with little choice other than to emulate those states that have proven successful in international affairs. As such, states will eventually all function and behave in essentially the same fashion, and a balance of power arrangement soon emerges. For such a system to emerge, asserts Waltz, all that is necessary is an anarchical international system and member states that wish to survive. He further explains that with the term *balance* he is referring to the state of the system as a whole and not to any particular subset of states within the system. Thus, there might be a given level of disequilibrium between two or more powers in a given balance arrangement, but other factors (such as the presence of equally balanced superpowers, for example) may be keeping the system itself in equilibrium despite the existence of such discord. Additionally, he argues that states do not need to actively pursue a balance strategy, because it is the natural result of normal relations between states in the international system. This point was argued by earlier scholars such as Claude (1962), who distinguished between “manual” systems where balancing strategy are actively pursued and “semiautomatic” systems where such a strategy is not as consciously or aggressively pursued. Morgenthau (1978) goes even further, asserting essentially that the emergence of equilibrium is an inevitable outcome of international relations.

Polarity and Balance of Power

There are various ways in which capabilities may be divided or distributed in the international system, leading to different forms of balancing between one, two, three, or more major states in the system. The ways in which balances of power are formed make a difference, scholars contend, because some balance arrangements are more stable and more capable of maintaining peace than are others (Snyder, 1997). The difficulty here is that there is disagreement on which balance arrangement brings the greatest

stability. Waltz (1979) contends that stability comes from having the fewest possible great powers competing for power. Thus, a bipolar balance of power would be best. Other research, such as that of Singer, Bremer, and Stuckey (1972), asserts that in the 1800s, the incidence of war actually decreased as the number of great powers increased. This would seem to indicate that a multipolar system is more stable. Although they do concede that the opposite trend was true in the 20th century (most notably during the cold war), it is possible that this can be attributed to other factors such as the emergence of nuclear weapons, not to the innate stability of the bipolar system itself. Since the 1990s, researchers have debated the long-term prospects for peace in a unipolar system, where the dominant power may have primacy but not enjoy true hegemony. History has seen all three systems in play, including the multipolar system during the Concert of Europe, the bipolarity of the cold war, and the current unipolar arrangement that emerged after the Soviet bloc collapsed. These and other examples of each system are discussed later in the chapter.

Bandwagoning and Collective Security as Alternatives to Balancing

Running counter to the logic of balance of power is the phenomenon of bandwagoning. When a weaker state is faced with a potential threat from a predominant power, balance of power theory states that the natural tendency is to form an alliance with others that would rise to the level of the predominant state and keep it in check, thus preserving the equilibrium. However, history has shown that this is not always the case. On many occasions, these weaker states will ally with that predominant power, surmising that they can benefit from the additional security that the combined capabilities brings and that they will share in the spoils of any future military or other strategic gains. Within the literature, the structural realists perceive balancing to be the dominant behavior (Waltz, 1979) and perceive bandwagoning behavior to be the approach taken by states dissatisfied with the status quo that are willing to go to war to disrupt the equilibrium. Classical realists, meanwhile, are more apt to accept bandwagoning behavior by weaker states as a rational and logical strategy for preserving their own security. Among the prevailing avenues of discourse in this debate is the role that methodology plays in stacking the deck in favor of one or another perspective on the matter. This is a more formal, institutional approach wherein states enter into mutual agreement that an attack on any member state (or any aggression on the part of any state within the alliance) would be met with a collective response from all committed partners. There has been a sustained discourse between proponents of each approach in the literature, which is summarized well by Doyle (1994) and in other sources.

Origins and History of the Concept in Practice

Emergence of the Concept in Practice

Balance of power was first put into effect at the international level in Europe in the 17th century, in the aftermath of the Treaty of Westphalia in 1648. This peace formalized a system that had been emerging among the great powers of Europe throughout the early part of the century, including the recognition of state sovereignty in the international system. States, eager to find paths to secular peace after roughly a century of wars based on religious and ideological differences, gradually embraced the logic that underpins balance of power theory, and by the end of the 17th century, it was the overarching approach of European nation-states to international relations and foreign policy. Anderson (1993) and others note that balance of power was not unanimously supported at the time, however. Criticisms of the approach included those believing that a unified Christian world was necessary to confront the threat of Muslim states. These critical voices lamented the secular and amoral nature of the balance of power system, even if they were not the loudest voices.

During this time, Britain became quite adept at working this system to its advantage, routinely playing the role of balancer to keep hegemonic aspirations in check and to make itself a central power in determining what the balance would be in the European system. This is a role that Britain would continue to play throughout the 17th and 18th centuries and beyond.

The balance of power system withstood its first major test in the 1680s, as Austria and Britain rose to counterbalance France's predominant position in Europe, a position that was diminished by France's concurrent and relative decline, thanks to mistakes made by Louis XIV and others. What eventually emerged was the first major international coalition constructed to counter a dominant state's power. This Grand Alliance was based on the League of Augsburg and included (among others) Austria, Spain, Saxony, the Dutch Republic, and Britain. Britain became, in effect, the balancer in this coalition, essentially dictating treaty terms and working to keep the coalition together in the face of French power. It is asserted by some historians that Britain at this time realized that its true potential to influence European politics was in working to maintain a balance on the Continent that they could use to their own advantage (Sheehan, 1996).

Institutionalization of the Balance Concept

By the start of the 18th century, balance of power was an explicit part of the diplomatic equation for all major powers in Europe, and its practice entered what some have called a golden age (Morgenthau, 1978). In fact, it seemed to be the organizing principle of foreign relations

(Anderson, 1993), on which all other decisions were based. Many reasons have been given for the prevalence of balance of power theory during this time, perhaps the best known of these being Rosecrance (1963). Among other reasons, he attributes the centrality of the theory to the absence of strong nationalist tendencies in most states, the prohibitively expensive nature of all-out war as compared to the relatively limited resources of states, and the use of so-called secret diplomacy (including bribery and espionage) to gain knowledge of others' capabilities and intentions. War during this time was more limited in nature, using smaller forces and constricted to practical goals and objectives (rather than the grand religious and ideological campaigns that preceded and would follow this period). States saw themselves and each other as members of a fluid, cosmopolitan system in Europe, where every state was important and where alliances shifted on a regular basis to keep balance. It helped that during this time, the rulers of many member states were in fact of a nationality different than the state they were ruling. This served only to enhance the cosmopolitan and collective membership mentality of the system. Events during the 1700s, however, would illustrate that not all states were of equal importance, and that developments between so-called great powers that harmed smaller states but did not significantly upset the balance among the more powerful states were allowed to occur. Such was the case when Poland was partitioned in the 1770s by Russia and Prussia, and neither France nor Britain chose to intervene.

The French Threat to Balance and the Concert of Europe

The next major challenge to the balance of power system in Europe was the expansion of French power under Napoleon Bonaparte in the first decade of the 19th century. With the exception of Britain and Russia, most of Europe was either under French control, pressed into alliance with France, or too decimated by war to effectively challenge further French expansion. The French forces were finally turned back by Russia in 1812, causing a domino effect that led to the general defeat of France soon after. As Europe subsequently tried to restore balance to the system, there was a general belief that the fluid, semiformal alliances that characterized balance of power in the 18th century would not be enough. There must be developed, it was believed, a way to make the forthcoming balance more permanent and institutionalized. The result was the Concert of Europe, a multipolar balance of power between Austria-Hungary, Britain, Prussia, Russia, and France. Jervis (1985) has asserted that the dynamics and characteristics of the Concert of Europe, which emerged immediately out of (and in direct response to) the events leading to the Congress of Vienna in 1815, can also be seen in the aftermaths of World War I (1919) and World War II (1945). This postwar period, which stretched into midcentury, was

indeed marked by a more structured system than that of the previous century, marked by a multitude of formalized agreements and an abundance of international meetings at which the great powers were represented. The Concert period was also more stratified than that of previous iterations of balance of power theory. Second- and third-tier powers were not afforded the same rights as the great powers, and they were often not considered in managing the rivalries and competing interests among Austria, Britain, Prussia, Russia, and France. In matters of foreign policy and international relations, these managing powers extensively deferred to and considered the concerns of one another. The agreements reached were then handed down to the other actors in the system, who were largely relegated to accepting and adhering to the agreements without much question. Some scholars of the period have referred to these lesser states as buffer states that helped to preserve the peace between the great powers (see Craig & George, 1990). In addition to lacking the resources and capabilities of these great powers, the lesser states were often dependent on one or more of the great powers for economic health and military security.

The Concert of Europe fell into disarray by midcentury, starting with a wave of revolutionary movements in Europe in 1848. By the 1850s, most notably with the Crimean War in 1854, the great powers had begun waging war with one another, a reality that would continue until the 1870s. The balance of power system that was so meticulously crafted and so painstakingly executed for decades was lost. The wars that took place radically altered the political landscape of several states, since borders were redrawn and many smaller states were assimilated into larger ones. The new balance that eventually emerged from these great-power clashes bore little resemblance to that of the Concert of Europe but instead reflected the increased and more direct competition that existed among the great powers in the new international balance. Factors such as growing nationalism and the push of the Industrial Revolution served only to intensify the loose and tentative nature of the new arrangement.

Balance of Power in the 20th Century

By the 1890s, a new and more stable balance of power began to emerge. Unlike the multipolar system of the early 19th century, the emerging order was decidedly bipolar in form, characterized by two competing alliances of great powers in Europe that encouraged a level of zero-sum-game thinking that had not yet been experienced. Without buffer states and with no viable option for any state to remain neutral (i.e., outside the bipolar confrontation), even moderately small crises had the potential to affect the balance and bring the system to the brink of war. Although this prospect may have initially had the effect of keeping existing alliances in place and preventing relatively small situations from getting out of hand, the system itself served to breed distrust and apprehension between

the two sides and helped to elevate tensions to crises and crises eventually to war in 1914.

The aftermath of World War I revealed a world much different than the one that existed when the war had begun. Perhaps the most significant changes related to the receding regional balances of power that had persisted as independent entities for centuries and to the inclusion of non-European powers (most notably the United States) into the ranks of the great powers. After the war, these regional systems took a back seat to the global balance that had emerged when Britain, France, Russia, and the United States allied to counter the threat posed by Germany and Austria-Hungary. The divide was bipolar amongst the so-called great powers involved in the war, with the opposition between sides resting on reasons that were decidedly more nationalistic and ideological than in past regional systems. This changing dynamic had a negative impact on the flexibility of the balance of power system, since states were discouraged from remaining ideologically neutral and thus could not effectively claim a role as balancer.

As they did in 1815, the victorious powers sought to make more permanent the emerging balance and the wartime alliances that had thwarted German aggression in Europe in World War I. This effort met with only partial success. One reason why this effort ultimately fell short was a failed attempt at creating a permanent venue for international diplomacy, the League of Nations. Another factor contributing to the failure of the post-World War I balance of power was the inability of the victorious alliance to sufficiently weaken and pacify Germany. In addition to all this, international actors not central to the immediate postwar arrangement were discounted or ignored by the major powers, actors that would eventually come to influence the system in ways that were not anticipated.

Morgenthau (1978) asserts that it was not World War I that changed the existing balance of power arrangement but rather a number of other transformations in the international system that had the greatest impact. Perhaps the most important of these is what Morgenthau referred to as “nationalistic universalism,” wherein states worked to export or otherwise impose their own system of political, economic, and moral value systems on others in the system in a way that had not been seen in prior centuries. This practice continued throughout the 20th century to be sure and made for a system that was both less flexible and less grounded on shared political and ethical values than had been the case in the past. In this way, struggles between adversaries in the new balance took on a much more ideological tone, one in which there was little room for compromise.

Balance of Power in the Nuclear Age

These dynamics became more defined in the aftermath of World War II. The first global balance of power arrangement, forged only two decades before the outbreak of

World War II, failed to prevent the recurrence of German aggression in Europe. It also failed to fully account for the actions of actors that fell outside the scope of the traditional alliance system forged by the great powers after World War I, including those of Japan and China. The World War I-era system would experience a significant reorganization by the end of the war in 1945. In addition, trends that had become significant leading into the 20th century, including the prevalence of nationalism and ideological differences, would rise to become major factors in shaping the new system by midcentury. These and other emerging dynamics would contribute to the creation of a new balance of power unlike any that had existed in the past three centuries.

The states that emerged as the world’s great powers after World War II were, for the first time, not based in Europe. In fact, the traditional great powers of Europe became distinctly second-tier powers under the new arrangement, with only limited ability to shape or influence the balance of power between those actors in the first tier. In addition to this distinction, there were other characteristics of the new bipolar balance between the United States and the Soviet Union that made the new system different from any other to that point. For one thing, the differences between these two great powers were primarily ideological in nature and built on a foundation of what Morgenthau (1978) called “nationalistic universalism,” a trend in the system that had begun to emerge in the aftermath of World War I. Each side was convinced of the moral and political superiority of its own ideology and worldview and was equally convinced that the other represented a threat to its own existence. This instituted a system that left no room for compromise, and that discouraged the existence of neutral states or balancers among the world’s second- and third-tier powers. It also became a stark zero-sum-game atmosphere, wherein any gains by one side were perceived to cause a comparative loss for the other. Competition for third-party actors was intense, and the importance of these states was often elevated because of the zero-sum nature of great-power competition. Morgenthau discusses the expansion of primary global actors into “empty spaces,” making a distinction between such behavior among European powers in the 19th century and that of the United States and the Soviet Union during the cold war. Another significant difference in the new postwar balance was the fact that the two leading powers were quite significantly beyond any other states in the system, to the point that they were appropriately labeled *super powers* in the new balance. Put simply, no other actor in the system was perceived as having sufficient power and capability to challenge either the United States or the USSR.

Much of the reason for this reality was the emergence of a new dynamic into the international system, what might most appropriately be labeled as the most significant and consequential characteristic of the post-World War II international system: the existence of nuclear weapons

(Gareau, 1962). Some of the early scholarship on cold-war international relations, including that of Burns (1957), posited that nuclear deterrence would take the place of balance of power concerns in managing the anarchy of the international system. Indeed, it seemed possible that the presence of such weapons of mass destruction and the reality of mutually assured destruction that emerged as a result might impose an order over anarchy that was not previously possible on the international stage. Brodie (1946), for example, wrote at the time that the presence of nuclear weapons would change the nature of military power such that the goal would no longer be to win wars but rather to avoid them altogether. Though not a proponent of this view, Waltz (1981) does acknowledge that the nuclear threat effectively eliminated the prospect of war between the superpowers and contends that the possession of such weapons by more states may have spread the “nuclear peace” and lessened the instances of war in the system. Deudney (1993) argues that this argument implies the ability of nuclear weapons, in theory at least, to overcome anarchy and impose order on the international system, though this was not Waltz’s intended message. Others, however, have asserted that although the nuclear dynamic altered how balance of power worked, it nonetheless was still the overarching consideration for each of the superpowers and the other states within their corresponding blocs in the international system (Snyder, 1965). Bull (1977), for example, stated that balance of power no longer focused on amassing enough military force to acquire and hold physical territory, but rather centered on amassing enough of a “second-strike” capability to retaliate even in the wake of a nuclear attack.

This bipolar balance of power was built on deterrence, which depended (at least to some degree) on the subjective judgments of actors and their adversaries with regard to capabilities, intentions, and the resolve to carry out stated and implied threats. As such, some have labeled this more of a balance of terror than a balance of power (see Sheehan, 1996). In the cold-war arrangement, war avoidance became a primary objective, to a degree not seen under previous arrangements, and deterrence became the primary mechanism with which to achieve that goal.

Future Directions

The events of the 1980s marked the demise of the bipolar balance of power that had dominated international relations for half a century. In the resulting international environment, rife with uncertainty and seemingly rapid change, there are at least two areas of inquiry sure to garner a large amount of attention from balance of power scholars: the long-term viability and consequences of a unipolar system and the growing role of soft balancing in shaping international relations.

The Viability of a Unipolar System and the Rise to Prominence of Soft Balancing

A unipolar system eventually emerged from the collapse of the Soviet bloc and the end of the cold war, the first time in history that so much power was held by one state with global interests and a global reach. This system brought with it consequences that international actors to that point not been forced to confront, including significantly altered alliance choices and bargaining strategies, among others (Walt, 2009). There is no consensus, contend Walt and others, about whether the current period of unipolar balance will be harmful or beneficial over the long term. Although proponents of hegemonic stability theory (e.g., Keohane, 1984) believe that a single, hegemonic leader can benevolently impose a greater degree of peace and order on the international stage, others fear that a single, unchecked power can lead to a sense of imperialism and unilateralism. Evidence of both can be found since the survival of the United States as the lone global superpower. Critics point to the Bush Doctrine, which advocates unilateral action and the possibility of preemptive war, as evidence that a unipolar distribution of power can bring about conflict in the international system. In fact, some contend that the current unipolar arrangement is more illusion than reality and that the system is poised for a return to multipolarity (Layne, 1993).

However, states in the current distribution are not nearly powerful enough to effectively challenge the power and capabilities of the United States. It is also true that, with the spread of democracy and a capitalist economy to previously undemocratic and noncapitalist countries, states are much less likely to risk war over policy or ideological differences. So what can weaker states do to counter the predominance of a lone superpower? The answer may be soft balancing, which entails smaller states countering the hegemon (or, in the present setup, the state with primacy), using diplomatic, economic, and other nonmilitary means to counter the dominant power (Pape, 2005; Wohlforth & Brooks, 2005). Joffe (2002) and others contend that this type of behavior could very well be both an acknowledgment that no power (and presumably no coalition) could effectively match the United States’ military capabilities but that the behavior is intended to avert hegemony and bring the world to a multipolar arrangement by nonmilitary means. Much is yet to be determined to this end, but there is still much to observe, investigate, and research.

References and Further Readings

- Anderson, M. S. (1993). *The rise of modern diplomacy 1450 1919*. London: Longman.
- Brodie, B. (1946). *The absolute weapon*. New Haven, CT: Yale Institute of International Studies.
- Bull, H. (1977). *The anarchical society*. London: Macmillan.

- Burns, A. L. (1957). From balance to deterrence: A theoretical analysis. *World Politics*, 9(4), 494-529.
- Claude, I. (1962). *Power and international relations*. New York: Random House.
- Craig, G., & George, A. (1990). *Force and statecraft* (2nd ed.). Oxford, UK: Oxford University Press.
- Deudney, D. (1993). Dividing realism: Structural realism versus security materialism on nuclear security and proliferation. *Security Studies*, 2(3-4), 1556-1852.
- Doyle, M. (1994). Balancing power classically: An alternative to collective security? In G. Downs (Ed.), *Collective security beyond the cold war* (pp. 133-168). Ann Arbor: University of Michigan Press.
- Gareau, F. H. (Ed.). (1962). *The balance of power and nuclear deterrence*. Boston: Houghton Mifflin.
- Hopf, T. (1991). Polarity, the offense defense balance, and war. *American Political Science Review*, 85, 475-493.
- Ikenberry, G. J. (Ed.). (2002). *America unrivaled: The future of the balance of power*. Ithaca, NY: Cornell University Press.
- Jervis, R. (1985). From balance to concert: A study of international security cooperation. *World Politics*, 38, 58-79.
- Joffe, J. (2002). Defying history and theory: The United States as the "last superpower." In G. J. Ikenberry (Ed.), *America unrivaled: The future of the balance of power*. Ithaca, NY: Cornell University Press.
- Kaplan, M. (1957). *System and process in international politics*. New York: Wiley.
- Kaufman, S. J., Little, R., & Wohlforth, W. C. (Eds.). (2007). *The balance of power in world history*. New York: Macmillan.
- Keohane, R. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton, NJ: Princeton University Press.
- Layne, C. (1993). The unipolar illusion: Why new great powers will rise. *International Security*, 17(4), 5-51.
- Levy, J. S., & Thompson, W. R. (2005). Hegemonic threats and great power balancing in Europe, 1495-1999. *Security Studies*, 14(1), 1-30.
- Midlarsky, M. (1983). The balance of power as a just historical system. *Polity*, 16, 181-200.
- Morgenthau, H. (1948). *Politics among nations*. New York: Knopf.
- Morgenthau, H. (1978). *Politics among nations* (5th ed.). New York: Knopf.
- Pape, R. (2005). Soft balancing against the United States. *International Security*, 30(1), 7-45.
- Paul, T. V., Wirtz, J. J., & Fortmann, M. (Eds.). (2004). *Balance of power theory and practice in the 21st century*. Stanford, CA: Stanford University Press.
- Pelz, S. (1991). Changing international systems, the world balance of power and the United States 1776-1976. *Diplomatic History*, 15, 47-81.
- Rosecrance, R. (1963). *Action and reaction in world politics*. Boston: Little, Brown.
- Rosecrance, R. (2003). Is there a balance of power? In A. Vasquez & C. Elman (Eds.), *Realism and the balancing of power: A new debate* (chap. 11). Upper Saddle River, NJ: Prentice Hall.
- Sheehan, M. (1996). *The balance of power: History and theory*. New York: Routledge.
- Singer, J., Bremer, S., & Stuckey, J. (1972). Capability distribution uncertainty and major power war, 1820-1965. In B. Russett (Ed.), *Peace, war and numbers* (pp. 19-48). Beverly Hills, CA: Sage.
- Snyder, G. (1965). The balance of power and the balance of terror. In P. Seabury (Ed.), *Balance of power* (pp. 184-201). San Francisco: Chandler.
- Snyder, G. (1997). *Alliance politics*. Ithaca, NY: Cornell University Press.
- Vasquez, A., & Elman, C. (Eds.). (2003). *Realism and the balancing of power: A new debate*. Upper Saddle River, NJ: Prentice Hall.
- Walt, S. (2009). Alliances in a unipolar world. *World Politics*, 61(1), 86-120.
- Walt, S. M. (1985). Alliance formation and the balance of world power. *International Security*, 4, 3-43.
- Waltz, K. (1979). *Theory of international politics*. Reading, MA: Addison Wesley.
- Waltz, K. (1981). *The spread of nuclear weapons: More may be better* (Adelphi Paper No. 171). London: International Institute for Strategic Studies.
- Waltz, K. (1993). The emerging structure of international politics. *International Security*, 18(2), 44-79.
- Waltz, K. (2000, Spring). Globalization and American power. *National Interest*, 55-56.
- Wilkinson, D. (1999). Unipolarity without hegemony. *International Studies Review*, 1(2), 141-172.
- Wohlforth, W. C. (1999). The stability of a unipolar world. *International Security*, 21(1), 1-36.
- Wohlforth, W., & Brooks, S. (2005). International relations theory and the case against unilateralism. *Perspectives on Politics*, 3, 509-524.
- Wohlforth, W. C., Little, R., Kaufman, S. J., Kang, D., Jones, C. A., Tin Bor Hui, V., et al. (2007). Testing balance of power theory in world history. *European Journal of International Relations*, 13(2), 155-185.

44

DETERRENCE THEORY

SETH M. GOLDSTEIN

The Ohio State University

BRIAN R. KING

Muskingum University

RICHARD ARNOLD

Muskingum University

Would Adolf Hitler have shied away from invading Prague if Great Britain had possessed thermonuclear weapons? Would the consequences of even a limited war been too terrible for even Hitler to contemplate? Such questions are illustrative of the what-ifs that tormented cold war scholars and policymakers who remembered the apparent failure of deterrence after Munich, and it is only with an appreciation of this historic event—and the cold war that followed—that development of the study of deterrence can be understood. This chapter is not intended as a comprehensive review of these literatures—such an enterprise is not feasible given space constraints—but rather has the more modest goal of highlighting two important and long-standing debates in the scholarly literatures on deterrence. After an introduction to the basic concepts of deterrence, debates on rational deterrence theory and reputation acquisition are discussed as products of the methodological proclivities that the close linkage between theorists and practitioners encouraged. Though these debates will not be resolved within these pages, the authors hope their examinations will illuminate potentially fruitful avenues of inquiry for future deterrence scholars. This analysis also underscores the practical and scholarly dangers of methodological myopia and illuminates the benefits of methodological pluralism to future studies of deterrence.

What Is Deterrence?

Deterrence can be formally defined as “the use of threats by one party to convince another party to refrain from initiating some course of action” (Huth, 1999, p. 26). On an intuitive level, the logic is fairly simple: The best way a school bully can rule the playground is to convincingly persuade classmates into believing that he has the desire, strength, and will to hurt them if they challenge him. These qualities of desire, strength, and will are found in the academic literature as interests, capabilities, and resolve, respectively, while the ability to be convincingly persuasive is generally referred to as credibility. The core of this idea—preventing one’s opponents from attacking for fear of expected retaliation—is not a new one. Thinkers as far back as Thomas Hobbes and others have contended that shows of strength could instill enough fear in others to deter aggression. However, because deterrence was just one strategy among many before World War II, the idea achieved its maximum impact in the theory and policy communities as a consequence of the nuclear age.

The acquisition of the H-bomb by both superpowers in the early 1950s reframed the calculus of decision makers. Against this backdrop arose the fear of nuclear escalation between superpowers. Although decision makers were divided on the certainty that general (nuclear) war would result from the outbreak of a conventional war, all were

convinced that the possibility was at least potentially great. The first school of thought presumed that because any aggressive action by the enemy ran a great enough risk of triggering a nuclear response, initiating a war would be an irrational action. This perspective led to the development of what was termed *massive retaliation* in the Eisenhower White House as a deliberate strategy to deter aggression against the United States. This approach, part of Eisenhower's New Look strategy in the 1950s (King, 2005), directed that any military aggression against the United States would be met with a full retaliatory response, which could include the use of atomic (later nuclear) weapons.

However, as China and the Soviet Union grew their nuclear capabilities through the 1950s, many questioned the certainty of American retaliation on the Soviet Union either directly (direct deterrence) or in defense of allies (extended deterrence) (Morgan, 1977). Simply put, massive retaliation in the face of a nuclear counterresponse against the United States made the strategy seem much less of a credible threat. A second, more controversial school of thought consequently emerged as strategists considered the possibility that nuclear weapons could be successfully used in a limited fashion. This thinking suggested that thermonuclear weapons need not always be weapons of extinction. Targeting the adversary's military while retaining a sufficient supply of weapons to annihilate the adversary's cities constituted the newly emerging counterforce strategy. From this point forward, the possibility of small-scale or tactical nuclear combat had to be taken into consideration. This possibility of limited nuclear war introduced a degree of uncertainty into nuclear deterrence, becoming a central component of strategic thought.

Before getting too deeply into historical cases and the progression of policy over time, however, it makes sense to step back and consider the evolution of the literature on deterrence.

The Political and Conceptual Lineage of Deterrence Theory

Deterrence theory emerged as a popular and prescriptive theory of international relations in the 1940s and 1950s, though it had already been around in some form for far longer. This popularity can be attributed to a number of factors, not the least of which is the fact that it fit nearly perfectly within the framework of the emerging cold war, making sense of a rapidly changing international environment. It is also true that deterrence theory, with its emphasis on the structural distribution of military capabilities, was eagerly received by scholars who largely subscribed to the realist paradigm at that time. Finally, it aligned congruently with the policy-making approach taken by

American political and military leaders facing what they perceived to be an aggressive Soviet Union.

The Development of Deterrence Theory During the Cold War

Early theories of deterrence generally agree that an opponent is deterred when a contingent threat to use force prevents an attack from occurring. This in principle occurs when the costs of attack are perceived to exceed the benefits. According to Jervis (1979), deterrence theory appeared in three waves (see also Levy, 1988). Much early writing on deterrence during the so-called first wave came from scholars working either directly or indirectly with the U.S. military, most notably from the RAND Corporation. Bernard Brodie (1946), perhaps the most notable of these early scholar-strategists, contended that in the nuclear age, the primary purpose of military force must be not to win wars but to prevent them from occurring. Such ideas did not go unnoticed by political and military leaders in the early years of the cold war, which one could argue greatly influenced the development of strategies such as Eisenhower's massive retaliation approach. Still in their infancy, early formulations like Brodie's focused behaviorally—that is, simply on the use of threats to prevent war and ensure peace.

By the 1950s, a second wave of scholarship on deterrence had taken hold, and the idea that the prevention of attack was acquired through military capability soon became something of a conventional wisdom despite the sparseness of empirical evidence as to the theory's validity. This willing acceptance of such a theory without the requisite verification was likely a product of its fit with the widespread structural view of international relations and its congruence with the prevailing realist perspective of most scholars and policymakers of the time. Jervis (1979) further contends that deterrence theory may have revitalized the realist paradigm at a time when it faced a slew of criticisms concerning its intuitive nature and lack of specificity. Second-wave research widened the methodological scope of inquiry in ways that the first wave did not consider. The work of scholars such as Ellsberg (1960), Kahn (1960), and others used mathematical models and simulations to generate theory that influenced the development of nuclear policy in the United States. Though criticisms persisted over time concerning the optimistic assumptions of this perspective (including the expectations for societal recovery following a nuclear war), it held an influence that would last throughout most of the cold war.

Perhaps the scholar with the greatest long-term impact on the study of deterrence during this time was Thomas Schelling (1960, 1966), whose work brought game theory fully into the discussion of deterrence. The emphasis of these works on strategic cost-benefit analysis laid the conceptual groundwork for what would later be called rational

deterrence theory (RDT). Schelling's (1966) *Arms and Influence* theorized successful deterrence tactics like signaling and manipulation of one's opponent and had a significant influence on how subsequent scholars would organize their own research by focusing on strategies that would lead to the absence of military challenges or successful bargaining. His employment of formal models was his most enduring contribution to the study of deterrence. The depiction of human agents making choices led subsequent scholars to begin asking questions about the psychological implications of successful deterrence at the individual level. Schelling also asserts that any study of deterrence must take into account the role of compellence, since the two were inextricably linked. Looking at the scholarship that was to follow over the coming decades, the bulk of the research on deterrence did not take this approach. Some contemporary scholars of coercive diplomacy, however, successfully joined the two topics.

The third wave of deterrence scholarship thus sought to counter perceived weaknesses in the real-world application of the rational deterrence approach at both theoretical and methodological levels. Theoretically, RDT was criticized (George & Smoke, 1974), as were other basic (often untested) assumptions of the second-wave approach about the extent to which human decision makers were able and willing to perform the calculations suggested by the cost-benefit model of RDT (Jervis, Lebow, & Stein, 1985; Steinbrunner, 1974). These third-wave scholars contended that misperceptions and other decision-making peculiarities within a state could have a substantial impact on the outcome of a case (Jervis, 1976, 1989; Snyder & Diesing, 1977). Such works examined the role that probabilities and calculations played in deterrence, including the chances for failure when decision makers failed to properly integrate information or pay attention to probabilities. Third-wave scholars also considered how domestic audiences determined responses in crisis (Lebow, 1981) as well as how the bureaucracy would implement these decisions (Allison, 1971). These deviations from the type of objective rationality posited by RDT, such scholars suggested, meant that RDT did little to help understand when deterrent threats were likely to succeed or fail.

To a large degree, these conclusions stemmed from third-wave researchers bringing into the study of deterrence methods that emphasized theory testing and empiricism, rather than relying on intuition or deductive reasoning. This included the use of rigorous structured, focused comparative case studies (see George & Smoke, 1974) that emphasized bureaucratic and psychological process-tracing models and constituted a distinct move away from abstract theory building. Later research further developed statistical models to test the propositions of RDT. Quantitative scholars tried to edge closer to the perceptual operational measures of utility that third-wave case studies employed, developing more nuanced measures of the balance of military capabilities by shifting focus away

from the overall distribution of power to the power that a state could project regionally in any given crisis. This research also developed quantitative measures of interests and resolve, attempting to correlate these factors with bargaining outcome. Still others stressed the role of fairness and reciprocity in stabilizing relations between adversaries. The seeming progress of deterrence research over time as described previously obscured a debate that came into sharper focus toward the end of the cold war.

Post-Cold War and Post-September 11, 2001, Scholarship on Deterrence

In a post-Soviet, post-bipolar international system, many wondered what role nuclear deterrence would play in America's foreign diplomacy. Additionally, as the stable, pro-status quo superpower structure collapsed, so did the relative peace that such a structure had provided since the 1940s. Numerous so-called hot spots soon appeared, most involving rogue states, failed states, or localized problems such as ethnic cleansing. Some scholars, pundits, and practitioners began to wonder whether a strategy of deterrence could be successful against rogue states and ethnic violence or if the concept had simply outlived its practical usefulness.

In the literature, what Knopf (2008) labeled the *fourth wave of deterrence scholarship*, there was a noticeable shift in focus to address these developments. Some research centered on the utility of using a strategy of deterrence against smaller powers, or it centered on the use of both general and immediate extended deterrence in the absence of any superpower zero-sum game. The works of Powell (1989), Wagner (1991), and others examined how asymmetry and iterations of the security dilemma could increase a state's fears of being subject to an external first-strike advantage. A few scholars questioned the value of deterrence at all in a unipolar system that lacked the cold-war framework that had given it its policy genesis (Payne, 2001). Some, such as Payne, talked of the need for a redesigned and diminished role for deterrence, while others seemed to abandon it altogether. On the policy side, although deterrence did not disappear from the arsenal of decision makers, it certainly became less prevalent. In its place, it seemed, were interventionist strategies and a greater focus on compellence.

On September 11, 2001, the global landscape changed again, and with it came another shift in focus for the deterrence literature (Knopf, 2008). Suddenly, it seemed, there was an abundance of research examining the utility and limitations of deterrence against terrorist organizations (Davis & Jenkins, 2002; Kenyon & Simpson, 2004; Lebovic, 2007; Trager & Zagorcheva, 2005), rogue states (Smith, 2006), and other nonstate actors on the world stage (Auerswald, 2006). This renewed examination of U.S. national security threats has led some to contend that perhaps deterrence is no longer a feasible strategy in the

21st century. Freedman (2004) asserts that the theory had already become unwieldy, the theory is disguised as a science, and in the face of terrorist threats some of the very fundamental elements of the approach are flawed. For example, it is not possible to use a strategy that depends in part on calculations of risk when the adversary doesn't even take risk into consideration. Indeed, it may be problematic merely to determine the proper audience for a deterrent threat, considering the clandestine nature of terrorist groups. He further argues that as interventionist policy in the 1990s was replaced by unilateralism and preemptive war in the 2000s, the nation's leadership had already discarded deterrence as a viable tool of foreign policy, a sentiment shared by many researchers in the field.

Throughout the post-cold war period, there have been numerous comprehensive reviews and reassessments of the theory and the practice of deterrence, as well as recommendations on how to rescue, salvage, and repair deterrence so that it might take a more modest place among other tools of the policymakers (Freedman, 2004; Zagare & Kilgour, 2000). This new wave of dynamic introspection and assessment does not seem to be slowing even a decade beyond the events of September 11, 2001, and more than two decades since the fall of communism in Europe (see Paul, Morgan, & Wirtz, 2009).

Enduring Controversies in the Study of Deterrence

Rational Deterrence Theory: Methodological and Epistemological Divides

In many ways a perennial problem of social science, the epistemological question concerns how we observe and gear our methodology to theoretically account for non-events. More generally, these issues pertain to how we interpret evidence. The problem: Deterrence is not noticed when it is working, but it is extremely conspicuous in its absence. Consequently, scholars spend a great deal of time explaining why the world went to war in 1914 but pay less attention to the fact that war was avoided for the preceding 45 years. As such, do events such as the Cuban missile crisis constitute a success for RDT? Do we see confirming evidence of RDT in the Soviet withdrawal of their missiles or disconfirmation of the theory in its inability to predict their initial placement (Jervis, 1989)? This picture becomes even more complex when one considers how many possible crises simply did not emerge because policymakers feared nuclear war or whether policymakers created self-fulfilling prophecies by studying the scholars of the day. This chapter focuses on two implications that spring from the assumption that mental deliberations can be nonproblematically inferred from behavior.

The first is the implication of many early quantitative works on general deterrence—that the sole variable

determining whether a state will go to war is the presence or absence of a credible threat. Some scholars assume states to be rapacious entities just waiting for the opportunity to expand. Statistical approaches to rational deterrence provide evidence that deterrence theory has succeeded in the main, especially in the post-World War II world by showing a correlation between structural predictors of successful deterrence and the propensity of disputants to back down in crises. Perhaps the best known works in this vein come from Huth and Russett. Huth and Russett (1984) used statistical analysis to argue that credible threats determined outcomes in immediate deterrent crises. Similarly, Reiter (1995) argued that RDT explains the relative absence of preemptive wars since 1816. The results of these statistical analyses are problematic, for their assumption that the correlation of nonevents (peace) with interests, capabilities, and resolve signifies success for RDT. The controversial logic that pervades studies of immediate deterrence is that states defined as challengers may have had no interest in changing the status quo and considered a challenge purely as a tactic of diversion or bluff, thus overdetermining their backing down in a crisis. Though more complicated models have attempted to account for the inferential problem by doing a better job controlling for the domestic and perceptual sources of deterrence, early quantitative research surely overestimated the frequency of RDT's success.

The case study literature, however, persuasively argues that RDT fails because challenges of general deterrence do occur in spite of credible threats, just as immediate deterrent crises sometimes end in war. George and Smoke's (1974) 11 case studies "demonstrate [a] sequential, gradual failure of deterrence" (p. 103). One cannot infer a success for RDT, George and Smoke argue, from the absence of conflict. Other process-tracing case studies (see Jervis, Lebow, & Stein, 1985) also provide powerful evidence that RDT is invalid not because deterrent threats are not credible but rather because these threats are rarely accurately perceived by decision makers.

But just as findings from the statistical tests of deterrence theory were predetermined by the assumptions implicit in the way the methodology interprets evidence, so too were the results of research employing the case study method. The near exclusive focus on crises in these works made certain conclusions more likely than others, especially those concerning the effects of psychological biases in causing failure of the rational deterrence model (see Jervis, 1976). Case studies usually focus on deterrence theory in crises because assessing the validity of RDT in noncrisis cases is problematic given the lack of archival materials on nonevents like the absence of war. But because crises are the circumstances in which RDT is most likely to fail, case studies are thus prone to sampling issues by selecting on the dependent variable—in this case the breakdown in general deterrence that leads to a crisis (see King, Keohane, & Verba, 1994)—a practice that can

overstate RDT's inability to predict outcomes. Similarly, reliance of case studies on political memoirs can indeed be problematic because they are written after crises and are not necessarily indicative of actual deliberations as a crisis played out. Further, critics contend, because RDT is concerned only with choices, not the details of deliberations, evidence of misperceptions or failures to calculate probabilities in specific ways does not undermine RDT.

The point here is not that case studies cannot produce unbiased conclusions or that statistics cannot generate reasonable inferences. It is rather that there is great difficulty in assessing the validity of RDT, and the methods by which one does so often predetermine the focus of inquiry, the methodology employed, and thus the results. Assumptions made on the part of statistical analyses cannot help but do violence to the nuances of history in their efforts to achieve unit homogeneity. Conversely, comparative case studies may see the issue of bias in decision makers during crises as proof positive that RDT is wrong. But neither of these interpretations has a monopoly on wisdom. No method is without its biases or assumptions. What each method measures is as much a product of its own assumptions as it is an observation of positivistic truths about the validity of RDT. More honesty on this point would serve not only deterrence scholars but also the enterprise of social science as a whole.

Game Theory, Deterrence, and Reputation

We also see that unacknowledged assumptions of methodology play a role in the dispute over the importance that reputations for resolve had to deterrence credibility. Von Neumann and Morgenthau (1944), among others, provided the basic mathematical tools for game theory to develop, and immediately their work was seized on by politicians for its ability to prescribe best responses to Soviet actions and vice versa. It is from this dynamic—the interplay between mathematically inclined academics and military strategists (outlined in the section titled “The Political and Conceptual Lineage of Deterrence Theory”)—from which one of the enduring debates the cold war stemmed. Indeed, the concern over American reputation during the cold war, reflected in the conduct of the wars in Korea and Vietnam, may owe its existence to the intuitive plausibility of the assumptions of certain game theoretic metaphors for conflict. Before examining one of these assumptions, a bit of background is warranted.

Bertrand Russell may have been the first to liken the subset of deterrence that Secretary of State John Foster Dulles at the time called *brinkmanship* to a game called *chicken* (Schelling, 1966). The chicken game is widely acknowledged to be the most appropriate game theoretic representation of immediate deterrence—that is, deterrence in circumstances when an aggressor state is considering challenging the status quo (Snyder, 1971). General deterrence, on the other hand, because it is more dependent

on military capabilities and arms acquisitions, may be more accurately captured by the prisoner's-dilemma game. Because the most common policy debate was a question of whether crisis behavior affected general deterrent credibility through reputation creation, this chapter begins examination of the tenability of this proposition through examination of the hedonic sources of the chicken game's utilities.

In the chicken game, two teenagers arrange two cars a few hundred yards apart, facing each other on a one-lane road, with an audience watching the contest. In the classic formulation of the game, the cars then drive toward each other, with the object being, as the name of the game implies, to force the other driver off the road. The game is pitched as a test of courage, with each driver choosing to continue on or swerve at the last possible moment. A collision is the worst outcome for each driver personally and for the two drivers collectively. In this sense, the drivers want to avoid this outcome but not at all costs, because to swerve while one's counterpart does not results in the second-worst outcome for the individual who swerves. In this case, the driver who swerves while the other continues is the chicken. Each driver determines the appropriate move to make by assessing the other's likelihood of continuing on toward collision. This likelihood is calculated as is the credibility of a deterrent threat—that is, by assessing the other driver's interests in continuing, his or her capability to do so, but most importantly by gauging his or her resolve (or commitment) to doing so. And so in the chicken game, as in crisis deterrence, the contest is a test of wills that requires participants to ask themselves how willing they are to risk disaster in order to prevent the loss of face that would come from swerving. Convincing an opponent to swerve can be aided by convincing him or her that one has irrevocably decided to not swerve, perhaps by throwing the steering wheel out the window. This may up one's chances of compelling the opponent to swerve, but it also escalates the risk of collision if the opponent chooses to do the same.

What sorts of costs and benefits determine the utilities in each outcome of the game? The possibility of a collision produces tangible material harm—harm that would be apparent even to a visitor from another planet. In the case of swerving while the other continues (the chicken payoff), on the other hand, what each suffers is completely a product of what the opposing driver (and observers of the game) might think. In this sense, the damage done by cooperating while the other driver defects is not tangible harm but rather is simply the harm of losing face and looking cowardly in the face of a challenge. This sort of intangible harm may not be obvious to the naive observer, given his or her lack of familiarity with the primitive social psychology that teaches the participants of the game the negative consequences they will suffer if they swerve, now and in the future. The fear of being labeled a chicken is a fear of another's perception

and the possibility that one will get a reputation for swerving, not a fear of some material state of being. Implicit in the chicken game's utilities are embedded assumptions about the importance of reputation and the presence of a watchful audience who ascribes that reputation to the chicken—after all, why would being a chicken be bad if one cared not about what observers or an opposing player thought? In the game, players care about their reputation for resolve because their commitments are interdependent—what they do here will affect their ability to get what they want elsewhere. This underappreciation of the importance of the source of utility within the chicken game metaphor laid critical groundwork for subsequent controversy among deterrence scholars and practitioners about whether (and if so, what types of) reputation matters to credibility.

*Reputations for What Characteristics
and for What Actors?*

Schelling (1966) famously argued that reputations for resolve mattered because American commitments were interdependent. Accordingly, behavior in a peripheral crisis in which American interests were not obviously engaged would affect the perceived strength of American commitments in areas of vital interest. “[The United States] lost thirty thousand dead in Korea to save face for the United States and the United Nations, not to save South Korea for the South Koreans, and it was undoubtedly worth it” (pp. 124–125). There is a certain amount of intuition behind the idea that reputations of some sort should matter, but these intuitions do not constitute proof that reputations for resolve matter to deterrence.

Most people readily acknowledge that the external validity of any abstract model is affected by the semblance its structure and assumptions bear to empirical reality. By the chicken game's logic, decision makers in a crisis should make choices based on the premise that other actors will use present crisis behavior to assess future general and immediate deterrence credibility. No policymaker, however, could know that reputations for resolve mattered; they simply assumed that they did. Indeed, the whole of nuclear deterrence was and is a speculative enterprise by nature; because no one had any experience fighting thermonuclear wars, no one knew exactly what prevented them. Though it seems clear that policymakers worry about their reputations for resolve, the empirical evidence that this concern is warranted is far less convincing. Empirical work on reputations for resolve grew up in the shadow of the cold war and has converged around the idea that although reputations can matter to crisis credibility, the types of reputations that matter are not the ones that policymakers during the cold war most predominantly feared.

Seminal efforts to empirically assess the importance of reputations for resolve included Hopf's (1994) study of the

inferences Soviet decision makers drew from American losses in peripheral areas of the globe. If reputations for resolve formed based on American behavior in the periphery, one would expect American losses, withdrawals, or both to lead the Soviet Union to conclude that U.S. core commitments to defend their vital interests were incredible on this basis. Hopf instead found that American retreats were explained away as a result of biased inferences that led Soviet decision makers to not focus on an American loss of face but instead on how the United States would never again allow such a loss in the future, especially in a strategically critical region. Furthermore, aggressive peripheral American interventions solidified Soviet perceptions of an imperialist America and thus brought about the very Soviet challenges they were designed to prevent. Though Crescenzi (2007) found evidence that extradyadic reputations correlated to the likelihood of extradyadic general deterrent challenges, the type of reputations that mattered were not reputations for resolve as assumed in the chicken game. Rather, Crescenzi's conclusion concerned “reputations for hostility,” and so his results are complimentary of Hopf's in suggesting that peripheral foreign policy interventions lead to reputations for hostility and thus are frequently counterproductive.

Mercer (1996) uses attribution theory as an empirical device for the study of reputations for resolve, arguing that a reputation can only form when a dispositional, rather than a situational, inference is drawn about behavior that may be classified as irresolute. Observers' attributions were dependent on whether they were explaining the behavior of an ally or an adversary and the desirability of said behavior. Generally, Mercer contended that desirable behavior was attributed situationally whereas negative behavior was explained dispositionally. By this logic, adversaries can easily get reputations for having resolve, since undesirable behavior is explained dispositionally, but do not get reputations for lacking resolve since desirable behavior was explained away situationally by virtue of its desirability. Contrarily, allies can easily get reputations for lacking resolve but rarely for having it. Press (2005) went even further, suggesting not only that reputations for irresolute behavior don't form, but also that decision makers do not use another state's past actions in any way to appraise their credibility. Rather, they rely instead on the “current calculus” of interests and capabilities to determine credibility.

Studies of many methodological stripes would disagree with these sweeping claims. Quantitative analyses have suggested that irresolute behavior informed calculations of credibility but only in cases involving repeated intradyadic crises (Huth & Russett, 1988). Although the inferences that policymakers feared would be drawn as a result of irresolute behaviors in the periphery were not drawn, behavioral evidence indicates dyadic reputations for a lack of resolve did seem to form. So although the research program on reputation indicates that peripheral interventions

to uphold reputations for resolve may be inadvisable, it does suggest that intradyadic behavior now can affect threat credibility later because other types of reputations do form. A great deal of work suggests that the dynamics of enduring rivalries offers microlevel reasons to buy into this conclusion. Enduring interstate rivalries persist because conflicts can spiral recursively (Jervis, 1976), and decision-makers' interpretations of any novel situation are held prisoner by their prior beliefs (Tetlock, 1999). Enduring dyadic rivalries thus shape the images actors hold of each other and can affect the ability of states to make successful deterrent threats as well. Regardless of the movement toward scholarly consensus, the intuitive plausibility of the idea that U.S. behavior in any given conflict affects its credibility in future crises ensures that the debate will rage on in domestic political forums in the coming years. Examination of this debate, however, does strongly suggest that U.S. efforts to manage and manipulate its reputation in peripheral military adventures during the cold war were, in fact, unnecessary.

Though most of the work cited previously assumes that reputations can accrue only to states, there have been increasing suggestions that leaders acquire reputations that affect coercive bargaining processes. Recent scholars have suggested that a leader's tenure and his proximity to his next election influence the reputations international actors attribute to them. As the fields of general and immediate deterrence coalesce into the unified field of bargaining, scholars have realized that reputations accrue to both leaders and states and that more work is needed to understand the way that multiple levels of analysis interactively generate coercive credibility in the international arena (see Greenstein, 2010).

Future Directions

The preceding section concluded that movement away from the state-state relations in deterrence inquiry may prove fruitful. This suggestion is perhaps reflective of the fact that U.S. security priorities are different today than they were during the cold war; the uncertainty of rogue states and substate terrorist groups have replaced the certainty of mutually assured destruction. Scholars have already begun work on these questions, with an abundance of research examining the utility and limitations of deterrence against terrorist organizations (Davis & Jenkins, 2002; Kenyon & Simpson, 2004; Lebovic, 2007; Trager & Zagorcheva, 2005), other nonstate actors (Auerswald, 2006), and rogue states (Smith, 2006). If this chapter has one critical conclusion, it is that deterrence scholars have too frequently lost sight of the assumptions and limitations of their theories and methodologies. Just as policymakers during the cold war learned from what they considered to be the failures of Munich, today's scholars must be vigilant in this new era to avoid repeating the mistakes of yesteryear. If the recent push toward more conscientious multimethod

approaches in political science is any indication, the early indications of lessons learned are fairly promising.

References and Further Readings

- Allison, G. (1971). *The essence of decision*. New York: Longman.
- Auerswald, D. (2006). Deterring nonstate WMD attacks. *Political Science Quarterly*, 121(4), 543–568.
- Brodie, B. (1946). *The absolute weapon: Atomic power and world order*. New York: Harcourt Brace.
- Crescenzi, M. (2007). Reputation and interstate conflict. *American Journal of Political Science*, 51(2), 382–396.
- Davis, P., & Jenkins, B. (2002). *Deterrence and influence in counterterrorism: A component in the war on al Qaeda*. Santa Monica, CA: RAND.
- Ellsberg, D. (1961). The crude analysis of strategic choices. *American Economic Review*, 15(2), 472–478.
- Freedman, L. (1989). *The evolution of nuclear strategy* (2nd ed.). New York: St. Martin's.
- Freedman, L. (2004). *Deterrence*. Cambridge, MA: Polity Press.
- George, A., & Smoke, R. (1974). *Deterrence in American foreign policy: Theory and practice*. New York: Columbia University Press.
- Goldstein, S. (2010). *The problem of leaders resolved: The effects of domestic dissent on coercive credibility*. Unpublished manuscript, The Ohio State University.
- Hopf, T. (1994). *Peripheral visions: Deterrence theory and American foreign policy in the third world, 1965–1990*. Ann Arbor: University of Michigan Press.
- Huth, P. (1999). Deterrence and international conflict: Empirical findings and theoretical debates. *Annual Review of Political Science*, 2, 25–48.
- Huth, P., & Russett, B. (1984). What makes deterrence work? Cases from 1900 to 1980. *World Politics*, 36, 496–526.
- Huth, P., & Russett, B. (1988). Deterrence failure and crisis escalation. *International Studies Quarterly*, 32(1), 29–46.
- Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.
- Jervis, R. (1979). Deterrence theory revisited. *World Politics*, 31(2), 289–324.
- Jervis, R. (1989). Rational deterrence: Theory and evidence. *World Politics*, 41(2), 183–207.
- Jervis, R., Lebow, R. N., & Stein, J. G. (Eds.). (1985). *Psychology and deterrence*. Baltimore: Johns Hopkins University Press.
- Kahn, H. (1960). *On thermonuclear war*. London: Oxford University Press.
- Kenyon, I., & Simpson, J. (Eds.). (2004). Deterrence and the new global security environment [Special issue]. *Contemporary Security Policy*, 25(1).
- King, B. (2005). Presidential management of international crises: *Structured management approaches and crisis learning*. Doctoral dissertation, University of Cincinnati. Available at <http://olc1.ohiolink.edu/search~S0/?searchtype=t&searcharg=Presidential+management+of+international+crises&SORT=D&extended=0&SUBMIT=Search&searchlimits=&searchorigarg=aking>
- King, G., Keohane, R., & Verba, S. (1994). *Designing social inquiry*. Princeton, NJ: Princeton University Press.
- Knopf, J. (2008). *The fourth wave in deterrence research: An appraisal*. Paper presented at the Annual Meeting of the American Political Science Association, Boston, MA.

- Lebovic, J. (2007). *Deterring international terrorism and rogue states: U.S. national security policy after 9/11*. London: Routledge.
- Lebow, R. N. (1981). *Between peace and war*. Baltimore: Johns Hopkins University Press.
- Levy, J. (1988). When do deterrent threats work? *British Journal of Political Science*, 18(4), 485–512.
- Mercer, J. (1996). *Reputation and international politics*. Ithaca, NY: Cornell University Press.
- Morgan, P. (1977). *Deterrence: A conceptual analysis*. Beverly Hills, CA: Sage.
- Paul, T. V., Morgan, P., & Wirtz, J. (2009). *Complex interdependence: Strategy in the global age*. Chicago: University of Chicago Press.
- Payne, K. (2001). *The fallacies of cold war deterrence and a new direction*. Lexington: University Press of Kentucky.
- Powell, R. (1989). The theoretical foundations of strategic nuclear deterrence. *Political Science Quarterly*, 100(1), 75–96.
- Press, D. (2005). *Calculating credibility: How leaders assess military threats*. Ithaca, NY: Cornell University Press.
- Reiter, D. (1995). Exploding the powder keg myth: Preemptive wars almost never happen. *International Security*, 20(2), 5–34.
- Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Schelling, T. C. (1966). *Arms and influence*. New Haven, CT: Yale University Press.
- Smith, D. (2006). *Deterring America: Rogue states and the proliferation of weapons of mass destruction*. Cambridge, UK: Cambridge University Press.
- Snyder, G. (1971). “Prisoner’s dilemma” and “chicken” models in international politics. *International Studies Quarterly*, 15(1), 66–103.
- Snyder, G., & Diesing, P. (1977). *Conflict among nations*. Princeton, NJ: Princeton University Press.
- Steinbrunner, J. (1974). Beyond rational deterrence. *World Politics*, 28, 223–245.
- Tetlock, P. (1999). Theory driven reasoning about plausible pasts and probable futures in world politics: Are we prisoners of our preconceptions? *American Journal of Political Science*, 43(2), 335–366.
- Trager, R., & Zagorcheva, D. (2005). Deterring terrorism: It can be done. *International Security*, 30(3), 87–123.
- Von Neumann, J., & Morganstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Wagner, R. H. (1991). Nuclear deterrence, counterforce strategies, and the incentive to strike first. *American Political Science Review*, 85(3), 727–749.
- Zagare, F., & Kilgour, M. (2000). *Perfect deterrence*. Cambridge, UK: Cambridge University Press.

45

RIVALRY, CONFLICT, AND INTERSTATE WAR

ERIC COX

Texas Christian University

The question of why states go to war has long been central to the study of international relations. The various answers put forward by scholars to this question have led to vigorous debates both between and within competing theoretical traditions that purport to explain state behavior. As this question has been studied over time, scholars have not settled on any one explanation for why some states choose to fight (or choose not to fight), why certain states enter into rivalries (or choose to end rivalries), or why factors that lead to conflict in one situation may not lead to conflict in another. For much of the post–World War II era, the study of conflict was greatly impacted by the ongoing cold war rivalry between the United States of America and the Soviet Union; this rivalry not only had the potential of leading to major war involving both of these countries and their allies but also had substantial impact throughout the world as new states were emerging from colonization and were often forced to choose sides in the larger cold war rivalry. The end of the cold war, however, led to numerous new questions, including how one state that, by most measures, was among the most powerful states ever to have existed could cease to exist almost overnight; whether a new rival to U.S. power would emerge; what shape conflict would take in the post–cold war era; and, perhaps most importantly, what the end of the cold war said about international relations scholarship, almost none of which predicted the end of the cold war.

This chapter on rivalry, conflict, and interstate war begins with an examination of the evolution of major theories of why states enter into rivalries and conflict. Given the expanse of explanations for state behavior, this chapter will focus primarily on realism and neorealism, the evolution of neorealism, constructivism, democratic peace theory, and rational choice explanations for the start of conflict. This section is not an exhaustive account of all theories of state behavior (see Chapter 37, “Realism and Neorealism”; Chapter 38, “Idealism and Liberalism”; Chapter 46, “The Democratic Peace”; Chapter 56, “Constructivism”; and Chapter 64, “Game Theory,” in this volume) but does capture much of the debate over why states fight. Next, the chapter turns to the ways in which empirical evidence is used to support or question various theoretical arguments. Finally, the chapter examines both real and potential policy implications of international relations scholarship before turning to a discussion of future avenues of research.

Theoretical Overview

Rivalry

The destruction wrought by World War I led to an immediate questioning of how the states of Europe had allowed themselves to enter into such a devastating

conflict and, as importantly, how to avoid future conflicts. Policymakers, led by U.S. President Woodrow Wilson, fervently believed in the principles of idealism, a view that appropriately designed international institutions that provided appropriate legal avenues to redress grievances, coupled with a system of collective security, could prevent major power wars in the future. The result of this belief was the creation of the League of Nations, an organization that would adjudicate disputes between states before they descended into conflict and, should one state attack another, lead a collective response of all member states against the aggressor state (collective security). The League of Nations faced many setbacks, not the least of which was the failure of the United States Senate to ratify the Treaty of Versailles, meaning that the United States, then the world's largest economy, would not be a member. As the League of Nations proved unable to successfully confront series of challenges in the 1930s, including German abrogation of the terms of the Treaty of Versailles, Italian aggression against Ethiopia, and the Japanese invasion of Manchuria, idealism as a both a theoretical and policy exercise appeared to be a failure. World War II simply furthered the point.

After World War II, Hans Morgenthau (1948) published his seminal work, *Politics Among Nations: The Struggle for Power and Peace*. In this work, Morgenthau argued that one of the primary failures of idealism was that it had attempted to impose morality on the state without recognizing the underlying cause of conflict: the perpetual pursuit of power. In the book, he argued that states did not abide by the same moral rules as individuals in their normal interactions; rather, each state was concerned primarily with the national interest, as defined by the pursuit of power. The most capable politicians would be those who appropriately considered each action in terms of whether it would benefit the national interest, not primarily whether it was moral. According to Morgenthau, war was the inevitable result of clashes between states competing to increase their power and could not be ended by the creation of international institutions, as the idealists had argued.

Morgenthau's (1948) work, though well respected, did have its flaws. Primary among them was that by arguing that states were always seeking power, his theory did not allow for variation that would explain when and why war was likely to occur. In particular, Kenneth Waltz (1979) argued that a good social science explanation had to have some variation in explanation (the independent variable) that explained a variation in outcome (the dependent variable). By not having any variation in explanation, Morgenthau's theory could not explain why states were not in a perpetual state of war. In summary, states seeking power was a background assumption, not an explanation for war. Waltz's ultimate answer to Morgenthau and traditional realism was *Theory of International Politics* in 1979, a work that he argued was

not an attempt to explain or predict a state's foreign policy behavior, but rather to provide a general explanation for the propensity for balances of power to emerge in the international system and, in general terms, what distribution of power in the international system is the most stable. In this work, Waltz laid the groundwork for the neorealist school of thought.

Waltz (1979) argues that the best explanation of state behavior is based on three assumptions: (1) States are unitary actors—each state can be thought of as an autonomous decision maker; (2) states are security seeking—each state pursues policies that will maximize its own external security, and (3) the international system is anarchic—there is no authority over states to which they must answer or to which they can appeal. Since states are driven primarily by security interests, states will seek to balance against states that are more powerful as no international agent could prevent the more powerful state from attacking them. If one state begins to emerge as more powerful than all others, alliances will shift to balance that state's power. Based on these assumptions, Waltz argues that the primary source of variation in the international system is the structure of the distribution of power between states. A system with one dominant state is unipolar, a system dominated by two states is bipolar, and a system dominated by more than two states is multipolar. According to Waltz, rivalries between states are driven by the logic of the balance of power. The Soviet Union and the United States competed with one another not because they were ideologically opposed, but rather because they were the two most powerful states in the system: The logic of the balance of power dictated that they be rivals.

Waltz (1979) went on to argue that war is less likely in a bipolar system than in a multipolar one. His argument rests on the difference between external and internal balancing. External balancing requires states to depend on allies in order to balance other powerful states in the international system; no state is powerful enough to balance others individually. Internal balancing, in contrast, is the process whereby a state uses its own resources to build up its security, not relying on allies. In a multipolar system, states have to rely on external balancing since no state is individually powerful enough to balance other states in the system; the competition for allies and the potential for one ally to pull others into a conflict makes external balancing relatively less stable than internal balancing. When states balance internally, as they are able to do in a bipolar world, they are less likely to compete for allies and less likely to be drawn into conflict due to the behavior of allies. Therefore, a bipolar system, according to Waltz, should be more stable.

Waltz's (1979) explanation of the distribution of power in the international system fit nicely with Robert Jervis's (1978) "Cooperation Under the Security Dilemma," an article examining the effects of anarchy in the international

system. Jervis expands on the concept of the security dilemma, which is related to the idea of balance of power. In essence, the security dilemma is the problem that in the anarchic international system, any action by a state to build up its ability to defend itself lessens the security of other states. The reason for this insecurity is that any increase in military capability of one state necessarily makes surrounding states less secure. In turn, those states build up their capabilities, which lessen the security of the original state, leading it to build up again. This cycle explains that both states with purely defensive intentions may find themselves in an arms race and, if one state determines that it must use its weapons because it cannot afford additional buildups, potential conflict. To this basic concept, Jervis added the variable of whether offensive weapons or defensive weapons were dominant in an era and whether this dominance could be perceived. In an era where most weapons and tactics are offensive in nature and this dominance is recognized by all, war becomes more likely as states must use their weapons offensively in order to defeat rivals. When defensive weapons and tactics are dominant and all states recognize this, war is less likely as states perceive attacking a rival to be more difficult. Like Waltz, Jervis's argument considers rivalry to be a function of anarchy: States are in security dilemmas with one another—rivalry relationships—because the nature of the international system is anarchic.

Other neorealists expanded on the theory or challenged some of Waltz's core assumptions. Stephen Walt (1987) expanded Waltz's original work, which applied only to great powers, to regional settings, arguing that state behavior was not necessarily always to balance against a larger power, but also, at times, to join, or bandwagon, with a more powerful state. His work focused on factors other than the balance of power to determine that states actually balance against threat, not just power. The factors that shape threat include power, geographic proximity, offensive power, and aggressive intentions. Geography is particularly important in a regional setting since geographically proximate states are more likely to have the military capability sufficient to attack one another and to compete over issues such as contested borders and resources. He argues that balancing is more common than bandwagoning behavior but that certain conditions favor the latter over the former. In particular, when states are faced with a much stronger power, they may bandwagon, particularly if no other ally is readily available. In addition, particularly aggressive states are more prone to provoke balancing behavior. For Walt, then, rivalry relationships are driven by a combination of factors—it is not only the power of surrounding states, but also their proximity and the perception that those states are threatening.

For critics of the neorealist explanation of rivalry and conflict, Walt's inclusion of aggressive intent revealed a fundamental flaw in neorealism: that it did not account for state intentions. In essence, by relying only on the security

dilemma and anarchy, neorealism had no real explanation for why states would choose to fight. As the cold war drew to an end and then ended spectacularly with the collapse of the Soviet Union, several different strands of scholarship rose up in response to neorealism. Neoclassical realism returned to the teachings of Morgenthau to conclude that power matters and states are security seekers but that, at its core, war is caused by states with aggressive intent, commonly referred to as revisionist states. These states are dissatisfied with the status quo in the world, particularly institutions that may be designed to maintain existing power structures, and seek to topple that existing order.

Other scholars argued that a focus on power alone was problematic if no attention was paid to the meaning of that power. Scholars in this new school of thought, known as constructivism, were led by Alexander Wendt (1992), particularly his article "Anarchy Is What States Make of It." In this article, Wendt argues that actors interact to create intersubjective meanings. In essence, states continuously interact to create and re-create their relationship. The identities they create affect their perceptions of the other's intent and the use of power. In this way, according to Wendt, U.S. military power will have different meanings for states depending on their relationship with the United States. Canada, for example, is far less concerned about American military power than Cuba. Similarly, the United States is far more concerned about one or two nuclear weapons in the hands of Iran than hundreds in the hands of Israel, or with China's nuclear arsenal as opposed to the United Kingdom's. If all that mattered were power, constructivists argue, the United States would feel just as threatened by the United Kingdom as China (or the Soviet Union in the cold war). For constructivist, then, rivalry relationships are created and re-created over time through interactions of actors. The United States–Soviet cold war was the result not just of the presence of nuclear weapons and large conventional forces, but also of the continual perception of the other as an enemy, and behaviors on both sides that reinforced that perception. Constructivist explanations allow for factors such as culture, religion, and national identity to play a role in establishing rivalries, though such factors are not always at play in constructivist work.

At the same time, for Wendt (1992), anarchy does not require states to be competitive or to enter into security dilemmas; rather, they are able to interact in other ways. Some states have managed to reorient their identities over time. For example, Germany and France have left rivalry behind in the context of the European Union; Egypt and Israel, once bitter combatants, have a long-lasting peace agreement; and Japan and the United States, bitter enemies in World War II, emerged as powerful allies.

Although many of these theories provide an explanation for why rivalries emerge, whether it is simply balancing behavior in the case of neorealism or the construction of rivalry identities in constructivism, often they provide less

explanation for the actual outbreak of war. To summarize, the theories may help explain which states are likely to fight but are less able to explain the actual initiation of conflict. This chapter now turns to a section that reviews some theoretical explanations for the initiation of conflict, though, like this section on rivalry, it is not a complete reflection on why wars occur.

War Initiation

Neorealism as described by Waltz (1979) and even as modified by Walt (1987) does not purport to explain the timing of decisions to go to war. Neither does constructivism; like neorealism, constructivism seeks to understand the reasons why rivalries emerge (and subside) between states but is not as effective at explaining the microcauses of war initiation. Several other veins of theoretical scholarship, however, purport to explain particular decisions to enter into conflict. This section focuses initially on variants of neorealist explanations for war before shifting to a discussion of expected utility theory.

Some neorealist scholars, drawing on the insight of scholars like Waltz, argue that much like rivalry, the decision to fight war is driven by the balance of power. In particular, many neorealists, particularly followers of power transition theory, coalesce around the idea that war is most likely to occur during a transition phase in the prevailing international power structure. In essence, when one state is beginning to surpass the strength of another, war is likely to occur. Scholars differ on which state is likely to initiate conflict; some argue that a rising revisionist state will challenge the old dominant power(s) as soon as it is able in order to confirm its status as the most powerful in the system and to overturn the existing order that was designed to benefit another state (see Organski & Kugler, 1980). Others argue that war is more likely to be initiated by the state in relative decline; as a rising state begins to challenge its place as the most powerful in the system, the old power(s) will initiate a war to prevent the new state from rising up (see Copeland, 2000).

Critics of this approach make the point that the focus on power alone still does not explain war initiation per se in that it does not purport to explain on which issues rising and declining powers will disagree. Constructivists, in particular, would argue that at the turn of the 20th century, the United States surpassed both the United Kingdom and Germany as the most powerful economy in the world and had begun to establish itself as a military power with the Spanish American War. Nonetheless, no one legitimately feared a British–American showdown. On the other hand, Germany and Britain did enter into a rivalry relationship that led eventually to war. This outcome begs the question of why one set of states fought while the others did not.

One large vein of literature that attempts to explain state decision making, including decisions of war and peace, is expected utility theory. Expected utility theory (EUT) is a use of rational choice theory to explain state

decisions. Though many scholars use variants of EUT, rational choice, or both, one of the most influential works leading to the theories development in international relations is Bruce Bueno de Mesquita and David Lalman's (1992) *War and Reason* (see also work by Fearon, 1995; Schultz, 2001). This work, like all rational choice theory, does rest on certain assumptions. First, these theories assume that individual actors behave as if they are rational. Second, actors have a set of preferences that they are able to rank order. Third, actors are able to assess the cost and benefits of different alternatives to achieve their preferences. Finally, by weighing costs and benefits, actors choose between these alternatives to maximize their expected utility. Bueno de Mesquita and Lalman focus on each state's leader as the actor ultimately responsible for making decisions regarding war and peace. It is important to note that neither they, nor most other rational choice theorists, argue that leaders are infallible. The leaders weigh different factors and make mistakes, either in the calculation of the costs of one action or the potential benefit of attaining certain desires. An actor's actions may also be influenced by norms and values—some actors may not consider certain actions because the actions are deemed unacceptable.

The theory works by looking at a given situation and determining what course of action helps the state (or leader) maximize utility. If two states disagree over a border, for example, the leaders of the two states must weigh different courses of action that are possible. For the purposes of this example, let us say that State A must choose whether to do nothing, initiate negotiations, or initiate an attempt to take the territory by force. To each action, State B has a similar set of choices (do nothing, negotiate, or respond with force). If State B does not respond with force to an initiation of force by A, it is presumed that A wins that interaction. One important aspect of EUT comes through in this example, however: For conflict to occur, the two states must have incompatible preferences. In this case, both states want the same piece of territory. In another case, it may be control over certain resources or influence in a third state. Power competition may matter in the calculations of leaders, but conflict occurs over conflicts in desired outcomes.

One possibility that EUT allows is that leaders will make decisions at least in part designed to keep them in power. The diversionary war theory, a variant of EUT, argues that leaders faced with political unrest at home may initiate a crisis or war in order to improve their domestic political standing by “gambling for resurrection” (see Downs & Rocke, 1994; Smith, 1996). Though not all theorists making this argument use an explicit rational choice framework, the argument fits well. Leaders weigh the possibility of losing office without initiating a conflict versus the possibility of losing office if they initiate conflict. If they determine that they are more likely to maintain office by initiating conflict, they may do so. Proponents of diversionary war arguments often point to the Falkland Islands

conflict between Argentina and the United Kingdom in 1982 as a prime example of leaders facing domestic difficulties (Argentina's in this case) initiating a crisis over islands possessed by the United Kingdom but claimed by Argentina, in order to improve their domestic standing. Argentina ultimately lost the conflict, and the military government in Argentina fell.

EUT does draw a number of criticisms. First, it assumes that state leaders do make rational calculations about the cost and benefits of various actions. Critics of this approach argue that most people are incapable of making such complex calculations and point to the numerous mistakes leaders seem to make in decisions to initiate conflict as an example. For example, Robert Jervis's (1976) *Perception and Misperception in International Politics* examines numerous causes of leaders making flawed judgments due to a variety of factors that impaired their decision making. In particular, leaders are often more accepting of information that confirms their previously held beliefs than information that contradicts it; leaders misinterpret the intentions of enemies, often seeing all actions by a perceived enemy in a negative light; and leaders may underestimate the cost of acting when they are predisposed to act and may overestimate costs when predisposed to do nothing. In response, rational choice theorists do acknowledge that people are not perfect decision makers but that, on balance, leaders exhibit what is termed *bounded rationality*. In essence, leaders do pursue the actions that they perceive to be the most beneficial, even when mistaken.

A second criticism is that rational leaders should probably never fight a war since some mutually agreeable negotiated settlement should exist that is preferable to suffering the cost of war. James Fearon's (1995) "Rationalist Explanations for War" answers this criticism by arguing that war is essentially an exercise in seeking information. If leaders knew the outcome of a war in advance, they would not fight because a negotiated settlement in which neither side would suffer the cost of war would be preferable. Leaders, however, lack certain information. They believe they know the relative military capabilities of the two sides, and they believe that they know their level of resolve—roughly the cost they are willing to bear to win (measured in lives lost, expenditures, time, etc.). Likewise, leaders have information about the other side's capability and resolve. These estimations, both of a state's own power and of the opponent's power, are imperfect. As wars are fought, information about these factors becomes available; when one side is clearly winning, the two sides update their knowledge about each other and are able to achieve a negotiated settlement. This explanation is also important in that it has been used by other scholars as a launching point to explain the reemergence of conflict and rivalries. Conflicts that end with clear victors are less likely to reemerge in the future than conflicts in which neither side is clearly defeated, since states have better information about their relative capabilities.

A third criticism of rational choice theory is more difficult to address: This criticism is that rational choice, while a useful tool, does not provide an effective explanation for preference formation of actors or for the methods that some actors are willing to use but others are not. Some rational choice theorists make room for the role of other theoretical explanations to establish the basic preferences of actors, including the possibility of normative restraints on options available to actors. More complex models take such factors into account by assigning them weights in a theory—some actors may weigh normative constraints more heavily than others. Some actors may be more willing to accept risk than others—something that can be modeled. The ability to include all such factors into a model does further the criticism, however, by suggesting that by incorporating so many possible factors into models, rational choice theorists may be providing insufficient explanations of behavior. In essence, in any particular case, a post hoc model can be constructed to provide a perfectly rational explanation for a state's or leader's behavior.

The best way to resolve this dilemma from a rational choice standpoint is to fully specify the source of an actor's preference. As noted previously, some scholars, including Bueno de Mesquita, assume that political survival is the ultimate goal of a leader. Decisions on rivalry, conflict, and war are made with an eye to how those decisions impact political survival for a leader or party. In this manner, decisions to go to war that could be detrimental to remaining in power—whether due to loss in war or domestic unpopularity—are unlikely to be made, while decisions that lead to a leader's staying in power will be made. In some instances, a leader may be unlikely to remain in power regardless of decisions made and may, in fact, gamble that a certain action, if successful, will result in his staying in power, no matter how unlikely that outcome is (a very small chance of staying in power may be better than no chance of staying in power). Even in specifying assumptions ahead of time, however, rational choice theories may be accused of post hoc rationalization—the idea that any action a leader made can be rationalized after the fact. By the same token, postevent analysis may also lead to the opposite criticism: that a leader should have realized the actions he was taking would lead to a fiasco.

In concluding this section, it is important to note that explanations for why states consider other parties to be rivals and why they fight may or may not be related. Neorealist theories may be combined to explain both (a) that states will balance against any state that is larger and (b) war occurs when one state is threatening to displace another in the power order. On the other hand, neoclassical realists, constructivists, and rational choice theories would all argue that there must be some form of disagreement for conflict to occur. For neoclassical realists, a revisionist state unhappy with the status quo may be necessary; for constructivists, two states must share concepts of one another

that preclude peaceful relations; for rational choice theorists, the explanation lies in incompatible preferences: Both states want something that requires a settlement with the other state. When one or both states believe it is more likely to achieve its preference through conflict, conflict occurs.

Empirical Evidence

Empirical evidence for the different theoretical perspectives is somewhat hard to adjudicate. In this case, the history of the world, even if the analysis is limited to the 20th century or post–World War II era, is so rich with examples that support for many different theoretical perspectives is possible. For example, neorealists examining World Wars I and II often disagree about whether Germany was a state on the rise threatening to supplant the United Kingdom or Germany was already the dominant state threatened by a rising Russia, then a rising Soviet Union (for the former view, see Organski & Kugler, 1980; for the latter view, see Copeland, 2000). Historical details can be argued to support both positions. Debates on empirical support for various theories in international relations often come down to debates on the appropriate methodology for studying the issue. Some scholars focus on qualitative approaches (e.g., Copeland; Mearsheimer, 2001), using in-depth analysis of a few select cases to demonstrate their arguments, while others use statistical analysis of interactions between many different states over time (e.g., Diehl & Goertz, 2000; Vasquez & Henahan, 2007). Scholars emphasizing the qualitative approach argue that studying cases in depth provides a more nuanced understanding of the factors that lead to certain outcomes. More quantitatively focused scholars argue that by looking only at a small subset of cases, qualitative work may not sufficiently explain a wide range of cases or may misapply lessons from a few exceptional cases. Quantitative work, on the other hand, is often criticized for being probabilistic—it can explain tendencies but may not explain the particulars of many cases.

Nonetheless, the combination of quantitative and qualitative work on rivalries has led to some interesting conclusions. First, interstate wars frequently recur between the same sets of actors, suggesting that rivalries between states are an important source of international conflicts. Although conflicts are often between similarly sized opponents (suggesting balancing behavior), that is not always the case. As important appears to be a factor suggested by Fearon's work and followed up by others: the importance of a clear victor. In wars where one side does not achieve a clear victory or outside actors impose a ceasefire, war is more likely to reemerge than if one state achieves its principal war aims. Second, wars are often about incompatible preferences. Several scholars, led by John Vasquez, have emphasized the importance of territorial disputes in predicting where wars will occur. In a similar vein, research on enduring rivalries—relationships between states marked by repeated conflict—suggests that rivalries are often likely to

emerge in postcolonial areas or after great power conflict, two times when borders are most in flux. A third finding in studies of enduring rivalries is that when one state in a rivalry is a democracy, the rivalry tends to be shorter, and if both states in a rivalry become democracies, it is likely to end. This finding is in keeping with the democratic peace literature that suggests democracies are highly unlikely to fight other democracies, though democracies may be just as war prone as other states with nondemocracies.

On this last point, recent studies have indicated that the interest of a state's leader may influence decisions on war and, at times, peace. New quantitative research indicates that, contrary to the diversionary war argument discussed above, leaders are most likely to start conflicts when they are secure in office. Although some examples of leaders gambling for their jobs may exist, on balance, most wars are started by leaders who are not facing significant opposition. This finding holds across regime types (Chiozza & Goemans, 2003).

Nonetheless, as continued disagreement on the real causes of World Wars I and II demonstrate, empirical questions regarding the sources of rivalries and conflict are not fully settled. International relations are quite complex; no simple, concise explanation has been found that explains the particulars of all conflicts or rivalries, though, as discussed previously, certain patterns have emerged from the data over time. Conclusions drawn from the data, however, may impact the policy decisions of states, as is discussed in the next section.

Policy Implications

Understanding the dynamics of rivalry formation and the initiation of conflict has the potential to help prevent future conflicts. Frequently, however, it seems that insights from international relations theory either do not inform the decisions of policymakers or are misapplied. This section focuses on aspects of U.S. decision making during the cold war, the U.S. war in Iraq, and U.S. actions vis-à-vis China.

The Cold War

During the cold war, the United States worked to maintain a system of alliances around the world to contain the Soviet Union. The United States fought wars on the Korean peninsula and in Vietnam to support allies that were not central to the defense of U.S. security in the interest of preventing the spread of communism. It also spent a great deal of money and provided arms to countries around the world in order to ensure that they were friendlier to the United States than to the Soviet Union, sometimes supporting insurgencies against Soviet-backed governments. Although this section focuses on the United States, it is important to note that the Soviet Union engaged in similar behavior. International relations theory in general would

not seem to be very sympathetic to these actions. Neorealists, particularly those influenced by Waltz, would argue that in a bipolar system, the primary focus of balancing behavior should be internal, not external. The United States may have expended resources it did not need to in order to confront the Soviet Union.

Neoclassical realism and constructivism would also seem to have suggested that the United States use a more cautious approach. For neoclassical realists, if the United States was perceived as a threat by engaging in aggressive acts, even ones intended to confront communism, balancing behavior against the United States should have been expected. Although U.S. policymakers seemed to have believed that the United States had to demonstrate its strength around the world to encourage states to ally with it, the predominant conclusion of neorealist theory is that states balance against the most powerful actors in the system. For constructivists, by continuing to treat the Soviet Union and all communist states as enemies and interpreting all actions through that lens, the United States helped to perpetuate the cold war dynamic.

Iraq

In terms of the war in Iraq, two prominent neorealist scholars, John Mearsheimer and Stephen Walt (2003), penned an article prior to the U.S. invasion of Iraq in 2003, arguing against the move. Part of President George W. Bush's motivation in invading Iraq was the spread of democracy to the Middle East, which seems to be based on the democratic peace theory that arose from international relations scholarship. Walt and Mearsheimer argued, however, that the war would not serve U.S. interest for several reasons, primarily because Iraq was not a true threat and could be easily deterred, contained, or both. They further noted that the war could have unintended consequences, and that an aggressive, unilateral foreign policy by the United States could provoke balancing behavior. Although the Bush administration did use the notion of democratic peace to help justify the invasion, on a larger level international relations theory seems not to have played a large role in the decision-making process.

China

Regarding China, different international relations theories suggest different futures—and recommended courses of action—for U.S. policy. If neorealism is correct and the natural tendency of states is to balance against more powerful ones, the United States should continue to expect China to build up its military and should, in fact, continue to arm itself in an effort to defend itself against China. In the neorealist view of the world as anarchic, states must be prepared to defend themselves. In this case, the United States must be ready to defend itself and its interests against China. Note that this does not mean that the United States should take an aggressive tone or provoke confrontations

with China—neorealists would not want the United States to hasten balancing behavior even if it is inevitable.

Neoclassical realism, although concerned about Chinese power, would also concern itself with Chinese intent. Is China a revisionist state that is dissatisfied with the status quo? If China is dissatisfied, then the United States should be alarmed. If it is not or if the United States can reasonably accommodate China without weakening its own position, the United States may have little to fear in a more powerful China. Finally, constructivism would suggest that the United States will have a major role to play in defining its future relationship with China. If the United States treats China primarily as a threat, China is likely to reciprocate, and the relationship may become one marked by rivalry, much like the cold war. If, on the other hand, the United States seeks to work with China as a partner, a rivalry relationship becomes much less likely. This debate on the preferred approach to China is ongoing both in government and among international relations scholars. Organizations such as the Council on Foreign Relations, a well-respected nonpartisan foreign policy organization, have held numerous forums and published several books examining this question, eliciting debate from individuals from multiple schools of thought.

To conclude, international relations theory often produces policy recommendations that are at odds with the policies ultimately chosen by the United States. Nonetheless, some prominent academics, including Henry Kissinger and Condoleezza Rice at the highest levels and individuals such as Joseph Nye and Larry Diamond at lower levels, have played a role in policy formation. Potential reasons that scholars of international relations have had a difficult time in influencing policy on questions of conflict may include (a) a focus on explaining past conflicts rather than predicting future behavior, (b) widespread disagreement among scholars on preferred policies, and (c) theoretic explanations that do not always readily translate into policy prescriptions. Finally, it may also be the case that theory could lead to self-fulfilling prophecies. If the United States engages in balancing behavior with China at the behest of neorealist scholarship and China reciprocates, both neorealists and constructivists could claim to be correct. Constructivists would be able to argue that the United States provoked the balancing behavior with its actions, while the neorealists would argue that China's behavior proved them to be correct all along.

Future Directions

The international system is in a constant state of change, a fact recognized even by scholars who argue that the basic motivations of states are consistent over time. Among the questions with which future scholars must grapple are these: (a) Will the United States and China emerge as rivals in a new cold war; (b) do terrorists organizations such as al Qaeda represent a true threat to the traditional

notions of statehood; (c) has economic globalization, particularly close ties between the economies of states like China and the United States, lessened the likelihood of future wars; (d) have normative shifts over the past 60 years made states less likely to resort to war; and finally (e) what will future conflict look like?

Each of these questions has an impact on the future study of international relations. Part of the answer to the first question may depend on the advice that international relations scholars provide to policymakers on the United States–China relationship. Question two addresses an underlying assumption of this chapter: that states matter. Future scholars must weigh the balance they give to interstate war and rivalry versus war between states and non-state actors. The third question asks whether economic ties and economic globalization have changed the way states interact. The United States is a major market for Chinese goods; any conflict between China and the United States would be bad from an economic standpoint. In answer to the fourth question, constructivists would argue that the use of war as an instrument of policy simply is not accepted among states as it once was and that states who engage in aggressive war—even the United States—face consequences for their actions. Finally, the weapons and tactics of war have evolved over time. Future wars may not look like past ones; for example, many states, now including the United States, have dedicated commands that focus on cyberwarfare. These are just some of the questions facing the field of international relations; they do not even address debates that still continue over the causes of past events or unforeseen circumstances that will arise in the future.

Conclusion

This chapter has provided an introduction to several schools of thought on what leads to rivalries between states and the causes of interstate war. The field of international relations has many schools of thought that attempt to answer these questions, and the schools of thought themselves are subject to great debates. While neorealists focus on the role of power and the natural tendency to balance against power due to the vagaries of international anarchy and the need for security, constructivists emphasize that states and leaders have a role in constructing their own reality. Although these theories may provide some insight into which states will be rivals (states competing for power or states who enter a cycle of competition), they may be less powerful in explaining the actual emergence of war. For that, theories that analyze the preferences and actions of leaders provide useful insights into the decisions that lead to war. Such arguments often focus on the incompatibility of preferences between actors. Finally, the empirical evidence available does not lead to definitive conclusions. Sadly, the reality of the persistence of war affords ample opportunity for theorists to argue over war's cause.

References and Further Readings

- Bueno de Mesquita, B., & Lalman, D. (1992). *War and reason: Domestic and international imperatives*. New Haven, CT: Yale University Press.
- Bueno de Mesquita, B., Smith, A., Siverson, R., & Morrow, J. (Eds.). (2003). *The logic of political survival*. Cambridge: MIT Press.
- Campbell, D. (1992). *Writing security: United States foreign policy and the politics of identity*. Minneapolis: University of Minnesota Press.
- Chiozza, G., & Goemans, H. E. (2003). Peace through insecurity: Tenure and international conflict. *Journal of Conflict Resolution*, 47(4), 443–467.
- Copeland, D. (2000). *The origins of major war*. Ithaca, NY: Cornell University Press.
- Council on Foreign Relations. (2002). *The rise of China*. New York: W. W. Norton.
- Diehl, P., & Goertz, G. (2000). *War and peace in international rivalry*. Ann Arbor: University of Michigan Press.
- Downs, G., & Rocke, D. M. (1994). Conflict, agency and gambling for resurrection: The principal agent problem goes to war. *American Journal of Political Science*, 38(2), 362–380.
- Fearon, J. (1995). Rationalist explanations for war. *International Organizations*, 49(3), 379–414.
- Gaubatz, K. (1999). *Elections and war: The electoral incentive in the democratic politics of war and peace*. Stanford, CA: Stanford University Press.
- Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.
- Jervis, R. (1978). Cooperation under the security dilemma. *World Politics*, 30(2), 186–214.
- Jervis, R., Lebow, R., & Stein, J. G. (Eds.). (1985). *Psychology and deterrence*. Baltimore, MD: Johns Hopkins University Press.
- Lemke, D. (2002). *Regions of war and peace*. Cambridge, UK: Cambridge University Press.
- Mearsheimer, J. (2001). *The tragedy of great power politics*. New York: W. W. Norton.
- Mearsheimer, J., & Walt, S. (2003, January–February). An unnecessary war. *Foreign Policy*, 134, 51–59.
- Morgenthau, H. (1948). *Politics among nations: The struggle for power and peace*. New York: A. A. Knopf.
- Organski, A. F. K., & Kugler, J. (1980). *The war ledger*. Chicago: University of Chicago Press.
- Schultz, K. (2001). *Democracy and coercive diplomacy*. New York: Cambridge University Press.
- Smith, A. (1996). Diversionary foreign policy in democratic systems. *International Studies Quarterly*, 40(1), 133–153.
- Vasquez, J., & Henehan, M. (2007). *Territory, war and peace: An empirical and theoretical analysis*. London, UK: Routledge.
- Walt, S. (1987). *The origins of alliances*. Ithaca, NY: Cornell University Press.
- Waltz, K. (1979). *Theory of international politics*. Reading, MA: Addison Wesley.
- Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 46(2), 391–425.
- Wendt, A. (1999). *Social theory of international politics*. Cambridge, UK: Cambridge University Press.
- Werner, S., & Yuen, A. (2005). Making and keeping peace. *International Organization*, 59(2), 261–292.

46

THE DEMOCRATIC PEACE

STEVEN R. LIEBEL

University of North Texas

War between nations is an ancient phenomenon. As long as there have been governments and groups with shared social and cultural identities, there has been conflict among them. A far more recent phenomenon, however, is the development and diffusion of democratic government. Democracy, as is recognizable today, was introduced through the revolutionary liberal movements of the United States and France in the late 18th century. Until the mid-20th century, the number of democracies in the international system remained markedly stable, increasing substantially only after the dissolution of the Axis powers and the end of World War II in 1945.

Concurrent with the spread of democracy, international conflict has become less frequent. A remarkable observation within this trend is that pairs of democratic states rarely, if ever, directly engage one another in violent warfare. This observed relationship between democratic states has come to be termed the *democratic peace* and is one of the most robust and influential findings in international relations scholarship.

This chapter reviews a selected literature linking democracy to conflict and peace among states. Although the nonwarring relationship between democracies is one of the most robust relationships in international relations, likewise, the study of democracies and war is substantial. Accordingly, cumulative knowledge is considerable. What follows within this chapter is separated into five sections. The chapter begins with a discussion of theories concerning

the international state system and the role democratic states play there. Second, empirical evidence and caveats within the theories are presented and investigated. Third, practical real-world applications of the democratic peace are discussed. Fourth, projections of future developments concerning the democratic peace are provided. Finally, this chapter concludes with a summary of the most important findings and issues.

Theory and Evolution

States and Sovereignty

It is impossible to examine the relationship between democracies and international conflict without first acknowledging the context within which democracies exist. One must first, then, identify the shell within which democracies exist, the state, and second identify the structural relationship shared by states in the international system.

Over the course of international relations scholarship, no single unit has been as influential or garnered as much attention as the state (Lake, 2008). States are territorial entities with defined political boundaries, controlled by an amalgam of institutions and rules and inhabited by a sizable population. State governments create and enforce policies to govern their territory. Governments send their soldiers to spill their blood in war and enforce the laws that

govern those same soldiers when they are at home. The central government of a state and its interaction with its population, then, is of primary concern of any study of interactions between states.

States are enabled by the existence of sovereignty. Sovereignty is most often characterized as the monopoly of legitimate use of force within a given territory. Sovereignty is thus an authoritative relationship between rulers and their population, in which political elites possess the legitimate right to enact policy and project force to carry the policy through (Lake, 2003). In addition to internal sovereignty exercised by local political elites, other states are not allowed to interfere with the inner workings of that state. There are thus internal and external components to sovereignty as a basis of legitimate monopoly of force. Sovereignty has laid forth the groundwork for what we currently know as a state and defines how states interact.

Variation Within States: What Makes a Democracy a Democracy?

When perceived as sets of institutions and rules of law protected by sovereign political boundaries, states are relatively easy to identify. What is not as simple in identifying, however, is the variation that exists among states. This classification is, however, necessary if one is to understand the behavior of said states on the international scene. By identifying differences between state institutions and subsequently differences in political practices within states because of those institutional differences, formal expectations of behavior can be produced. Domestic institutions, and the expectations that follow, are at the core of the democratic peace argument.

The most clear and useful distinction that can be made between states is that which differentiates democracy and autocracy. Democracies are states composed of electoral institutions where leaders are chosen by, and are beholden to, a domestic population, from among numerous contenders. This necessitates regular free and fair elections where the population is free to participate and no individual's vote counts for more than any others, providing the necessary impetus for leaders to adhere to their population's wishes (Dahl, 1998). Further, democracies have strict limits on the power of the government over their people. That is, citizens within democracies possess negative rights against their government, or a private sphere within which the government cannot interfere (e.g., Doyle, 1983). Leaders who pass repressive or otherwise unpopular policies that intrude into the population's private sphere can easily be removed from office through legitimate means sanctioned by law. Thus, democratic leaders adhere to limits imposed on them by their population, or they face certain punishment.

Contrary to democracies, autocracies generally do not possess electoral institutions. Neither are populations within

autocracies protected by a private sphere analogous to that which exists within democracies. Autocracies consist of very small, exclusive political regimes without credible elections and are often reliant on repressive policies to maintain office. Semiauthoritarian regimes exist between full democracy and autocracy. Semiauthoritarian regimes (also referred to as *anocracies*) exclude large portions of the population from political processes and can use moderately repressive means on occasion. This is not to say that elections cannot be held in either autocracies or semiauthoritarian regimes. However, autocratic elections are most often used for political cover to present an air of legitimacy to observant domestic and international audiences. Given the absence of institutionalized political practices, if an unpopular or invasive policy is passed, beyond removal from office by force or violence, it is very difficult to punish autocratic leadership (Goemans, 2000). As a consequence of these political structures, autocratic leaders are not beholden to their population and are largely free to do as they wish.

When attempting to empirically qualify what a democracy is, it is useful to conceptualize democracy as a three-dimensional cube. Contemporary scholarship ranks democracies and autocracies according to three scales: executive recruitment, executive restraints, and political participation (Gates, Hegre, Jones, & Strand, 2006). Each of these scales represents an axis of the cube. As any one of these categories increases in nature, the space occupied by the democratic cube increases, creating a more ideal democratic state. Likewise, as any of these categories decreases in nature, the cube's space is reduced, and the state approaches a more ideal autocracy.

The term *executive recruitment* refers to the ability of individuals to compete for the highest office in the country. The higher the percentage of a national population eligible to contend for political offices in the country, the more open the society to variation in popular leadership. The term *executive restraint* is concerned with formal limitations placed on the executive. The higher the level of restrictions placed on political elites, the more power is removed from the executive and is placed in the hands of the electing population. *Political participation* directly refers to the ability of the population to vote and voice its opinion on leadership free of threats, intimidation, and constraint. As the percentage of a population that can participate increases, and as that population owns more right of input into processes, leaders must speak to a broader audience and are thus accountable to a greater constituency. Participation also refers to the substance of political practice, that is, whether the population merely votes or if they have a hand in setting the agenda as well (e.g., Cohen, 1971).

Using the aforementioned criteria—executive recruitment, executive restraint, and participation—one can differentiate levels of democracy, anocracy, and autocracy. Ideal democracy is composed of regulated and competitive elections in which parties must be able to lose power, constrained executive authority, and openly competitive

participation through policy creation and voting by some large proportion of the general adult population. On the flip side, an ideal autocracy has closed elections in which only a select few individuals can run for office, unconstrained executive authority where political elites face little to no opposition, and intentionally constrained, non-competitive elections where a small proportion of the population can participate via voting, and that population has little to no say in setting the policy agenda. As the levels for these criteria rise and fall, it is easy to visualize the aforementioned democratic–autocratic cube increasing or decreasing in volume.

The institutionalized rules and practices of a state—that is, the respective level of democracy or autocracy—defines the manner in which the state and its people interact. To continue the present line of thought, the more ideal the democracy, the more open the society is to hearing and considering counteropinions and to contending in a legitimate manner with opposing perspectives; in essence, the more ideal the democracy, the more open a society is to nonviolent bargaining (e.g., Dixon, 1994). The mass population can express itself, and the government must listen. Heinous activity on behalf of leaders is punished through lawful removal from office via regular elections or impeachment. Likewise, unruly citizens can be punished through formalized judicial proceedings.

Whether the state in question is an autocracy or democracy, the domestic political environment permeates up to decisions made at the state level and is in turn represented to the world on the international scene. The domestic political environment assists in defining the parameters within which states interact internationally, and thus, the distinction between democracy and autocracy is critical to our understanding of international warring behavior.

Arguments of International Democratic Behavior

There are two prominent perspectives regarding the internal behavior of democracy and subsequently its impact on the international scene. This distinction must be made clear, since the behavior of democracies as they act alone and as they act in pairs are two separate and distinct theoretical persuasions. The first perspective concerns the behavior of democracies as stand-alone entities. This is the argument of democratic pacifism. Democratic pacifism is commonly referred to as the *monadic* theory of democratic behavior given its concern for a democracy acting individually and not in a relationship with any other particular state. The second perspective concerns the behavior of democracies interacting directly with one another. This argument is the focus of this chapter, the democratic peace. The democratic peace is a dyadic theory, as opposed to a monadic theory, because it is concerned with a pair (a dyad) of democratic states and how they act in accordance with one another.

Democratic Pacifism (Monadic)

Democratic pacifism argues that democracies are more peaceful in all of their international interactions than any other form of state in their relations with others. The argument that democracies are pacifistic and war averse has two primary branches: (1) the capitalist peace and (2) normative behavioralism.

The capitalist peace was championed by Schumpeter (1955) and has been advanced more recently by Gartzke (2007). According to this argument, a society based on open economic policies mitigates the desire for aggressive expansion by states. That is, war is generally seen as non-profitable given its interference with potential domestic economic gains. War is thus incompatible with the utmost preferences of states profiting from open capitalism and is avoided. Capitalist democracies were at the forefront of this line of thought and were considered the least likely to engage in violence with other nations. The population would have voice in rebuking interference in their practices and would pressure the government to avoid costly war. In states where the capitalist population benefits from their own nonviolent nature, war should be unlikely. It is worth pointing out that capitalist peace arguments do not argue that war could not be profitable and that capitalist states will never engage in violence with one another. Although at most a minimal probability at any given time, war is always at the least a minor possibility.

The second argument of democratic pacifism is concerned with the development of a deep liberal culture, which in turn minimizes the outwardly aggressive nature of democracies. The most important liberal ideals developed exclusively in democratic states are threefold: freedom from arbitrary authority, freedom to protect those rights already possessed, and the freedom to represent oneself through active participation in the government (Doyle, 1983).

The aforementioned rights develop in democracies a rich culture receptive to contrary opinions and bargaining (see, e.g., Dixon, 1994). Over time, individuals and groups within democracies are socialized into preferring nonviolent dispute resolution and mutual acceptance of contrary opinions. As a result, democracies are inherently more accepting of other states' ideological stances. In turn, democracies are less likely to villainize states that do not mirror the democracies' own political practices and ideologies. Ultimately, liberal democracies possess an inherent pacifism to which they alone are privy, and the resultant behavior is that democracies approach other states in a less contentious manner than autocracies, experiencing less war as a result.

The Democratic Peace (Dyadic)

The democratic peace is an argument that democracies, when paired together, rarely go to war with one another. To be clear, this does not mean that democracies never come

to low-level violent terms with one another. There are certainly instances where democracies engage one another in lower-level crises (Mitchell & Prins, 1999). Nor does it mean that democracies do not go to war regularly with autocratic states. In fact, democracies frequently engage in warfare with nondemocratic societies (Bueno de Mesquita & Lalman, 1992). Rather, the democratic peace argues that democracies, when paired together, are probabilistically highly unlikely to engage in violent warfare against one another relative to any other pairing of states (i.e., autocracy–autocracy or democracy–autocracy) (Bennett, 2006; Bremer, 1992; Maoz & Abdolali, 1989). There are three distinct theoretical perspectives as to what fuels the behavior of democratic pairings of states: (1) structural, (2) normative, and (3) informational.

Structural theory argues that the peace between democratic states is a derivative of the legal institutional structures of the democratic states. Democratic institutions bind state leaders and their policies to their national population. If a democratic leader sends his or her state into battle on grounds that the electorate finds insufficient, illogical, or damaging to themselves or the state, the leader can easily be removed from office via institutionalized rules of office: election or impeachment. As a result, democracies are war averse. This relationship, however, holds only when there are two institutionally constrained states (i.e., democracies) acting against one another.

The idea of a structural “perpetual peace” between democracies was first espoused in 1795 by Immanuel Kant (1795/1903) but did not gain significant attention until the early 1980s with the work of Michael W. Doyle (1983). Kant argued that states should base themselves on three “definitive articles.” These three articles are as follows: (1) the civil constitution of each state shall be republican; (2) the law of nations shall be founded on a federation of free states; and (3) the rights of men, as citizens of the world, shall be limited to the conditions of universal hospitality. Kant further argued that as an increasing number of states adopted the articles, a more pacifistic international union of states would develop, resulting in an international perpetual peace between states. Given the centrality of Kant to the democratic peace, these three articles are now discussed in further detail.

The first definitive article argues that a state’s preservation of basic individual rights, such as representation, is critical in determining the external behavior of a state. A representative republican state provides protection for the individual citizen to act autonomously from the state. Autonomous citizens, free of the reigns of an overarching government, are at liberty to voice opinions contrary to that of their state’s regime. Possessing the freedom to express an opinion vocally first, and institutionally through voting second, places a tremendous amount of power in the state’s population and beyond the grasp of political elites.

More important for Kant (1795/1903) than the ability of the people to voice dissent, however, was their will to

resist fighting a war not of their choosing. Ideally, according to Kant, a republican constitution embodies the ideals of those it represents. As such, the consent of those living under the constitution is required to determine whether there shall be a war. The citizens of the nation will be the ones spilling blood in war, not their leadership, and as a consequence it is the citizens’ prerogative to resist a burden unduly placed on them. By owning the right to representation and having limited barriers against domestic resistance to fighting, democracies should be less likely to engage in violent wars.

The second definitive article proposes that, over time, peaceful republics such as those discussed in the first article, shall join together under a common law in a pacific international union. This union of states, with each state sharing similar internal republican dynamics, will help individual states come together internationally to overcome and avoid violent tendencies through international acceptance and application of republican law. This international covenant of peace equates roughly to a single government composed of numerous nations and will expand only as the number of individual republics increases, reducing war between nations at the same time.

The third definitive article focuses on universal hospitality between states. Universal hospitality implies that no state or representatives of a foreign state shall be treated with hostility when they enter another’s country. Specifically, scholars have taken this to mean that no state shall impinge on another’s ability to trade commercially with another: States have “the obligation of maintaining the opportunity for citizens to exchange goods and ideas, without imposing the obligation to trade” (Doyle, 1983, p. 227). Commerce and trade between nations, it seems, is important to maintaining the collegial international union sought in the second definitive article (Russett & Oneal, 2001). However, evidence supporting the pacific effect of trade is mixed. In fact, the impact of trade on proclivity for violence is dependent on the volume of trade, trade dependency, and the form of good being traded (e.g., Barbieri, 1996; Crescenzi, 2003). In some instances, strategic advantage is assumed by a predatory trading partner increasing the chances of war. In most all cases, however, the impact of trade is mollified by either the presence or absence of democracy (Beck, 2001). The Kantian peace hinges much more critically on democracy than on trade. However, when the three definitive articles are adopted simultaneously, the evidence of dyadic democratic peace is substantial.

Moving beyond Kant’s theory of perpetual peace, the second theory of democratic peace, normative theory, claims that ritualized behavior within democracies creates what are often referred to as “norms of bounded contestation” between democracies (Dixon, 1994). Norms are expectations shared between states as to what constitutes proper and expected behavior. A democratic state behaves in a certain manner given its domestic environment. Over

time, the domestic milieu becomes entrenched and creates a solidified internal democratic behavioral culture. When acting in pairs internationally, democracies recognize the constraints placed on one another by their domestic environments and do not expect one another to violate their respective domestic obligations. These normative expectations are further instilled between states through their repeated interactions with one another. Ultimately, recognition and acknowledgment of these norms reduce paranoia brought about by strategic uncertainty between states, limiting potentially extreme behavior and in turn making democratic dyads less war prone.

The third argument of the democratic peace is based on information conveyance. Being transparent organisms, democracies are able to signal their intentions and capabilities more credibly than autocracies (Fearon, 1994). International observers can easily witness policy processes and any potential military activity. In doing so, democratic transparency reduces uncertainty between nations and ameliorates the need for preemptive violence.

This argument is contingent on the policy preferences of leaders and their electorate. Given the visibility of policy-making processes within democracies, when democratic leaders make decisions, domestic observers “assess the performance of leaders” (Fearon, 1994, p. 577) and punish or reward said leaders accordingly. The same inherent transparency of the democratic structure allows the state to credibly convey their intent internationally. A constituency supportive of aggressive behavior advocated by political decision makers is an effective signal to outside states that the democratic leader will act in a likewise aggressive fashion. Backing down in the face of domestic pressure would lead to the potential loss of office. The same can be said of nonaggressive policy; if both leaders and the electorate favor nonconfrontational means, the signal is easy to interpret internationally. For the leader to act aggressively in this situation would result in an unfavorable public reaction. Thus, the ability of a democracy to signal interests and intent should reduce observing states’ uncertainty about the democracy’s actions and should relieve paranoid onlookers of their inclination to preemptively act on their fears.

Empirical Evidence and Counterarguments

The democratic peace is a theory of human and state behavior. Humans and states are able to alter their behavior under any given circumstances. As such, the democratic peace is a probabilistic theory of behavior and not a concrete law governed by, for example, Newtonian laws of physics (as are billiard balls on a table). However, the empirical fact that democracies have so rarely gone to war has compelled some to deem the democratic peace a law or something very close to it. Levy (1988) was the first to accord democracies with this impressive trait: “The absence

of war between democracies comes as close as anything we have to an empirical law in international relations” (p. 662). More recently, Ray (1995) has argued in support of this claim, in that when definitions of democracy and war are sufficiently strict, there are no observed instances of wars between democracies.

Although the statistical absence of war between democracies is impossible to ignore, there are, however, five focal counterarguments to the democratic peace. First, one must confront definitional aspects of the theory. That is, how does one define *war* and *democracy*, and in turn, how do varying definitions alter the cases that support or violate the democratic peace? Second, states vary in respective levels of development. As such, one must question the validity of the democratic peace across both developed and developing democratic states and assess if the democratic peace holds across all levels of state development. Third, the theory of democratic pacifism must be investigated. That is, are democracies truly peace loving and risk averse as arguments of monadic democratic behavior would have us think, or are democracies wolves in sheep’s clothing? Fourth, the international context and its impact on democracies and war must be addressed. Finally, the dominant theoretical persuasion of the latter half of the 20th century, realism, will be discussed in relation to the notion that democracies, of all states, are war averse.

The Debate Over Definitions

Although theories diverge in their explanations of the democratic peace, the ultimate litmus test for any theory of behavior is the historical record. In this regard, evidence is overwhelmingly in support of an international peace between democracies. Of the 80 recognized interstate wars that have occurred between 1816 and 1997 (Sarkees, 2000), there have been fewer than a handful of wars involving democracies on contending sides. However, when one presumes a specific empirical relationship such as war between democracies, the relationship itself is dependent on how one defines both *democracy* and *war*.

Empirically, war has most often been defined as an occurrence of violence between two or more states involving at least 1,000 battlefield deaths (combatants, not civilians) per year between belligerents (Sarkees, 2000). This definition may seem arbitrary, but it is a fair barometer of the level of violence between parties (e.g., Ray, 1995). A cutoff of 25 battlefield deaths would qualify small riots and uprisings as wars. If the level were such, the number of wars in the international system would be incredibly high. Thus, the 1,000 battlefield deaths per year mark is sufficiently high as to exclude smaller-scale incidents such as localized uprisings, while simultaneously including those incidents that vary from the largest of international confrontations, for example, World War I and World War II (approximately 8 million and 16.6 million relative combatant casualties), to brief but intense wars such as the 4-day 1969

Football War between El Salvador and Honduras (1,900 casualties) or the 2-week 1973 Yom Kippur War between Israel and its Arab neighbors (16,401 casualties).

Although the aforementioned definition of war has been widely accepted by the scholarly community, the definition of democracy is far more ambiguous. As previously discussed, democracies are an amalgam of popular participation, executive restraints, and free and fair elections. Any study attempting to ascertain the proclivity for democracies to go to war must address the questions of what necessary levels of popular participation and executive limitations on power are. According to Ray (1995), in conjunction with one's definition of *war*, the definition one uses for *democracy* alters the number of wars that could potentially violate the democratic peace dramatically.

In the strictest sense of democracy—a system with regular elections for political leaders, at least 50% of the adult population votes or is eligible to vote, there is more than one political party within the government, and governing parties can lose elections—and using the 1,000 battlefield deaths marker as the definition of war, there are, in fact, no wars between approximately 1800 and 1997 that would classify as wars between democracies (Ray, 1995). However, there are a number of conflicts between states approximating democracy or of such intensity as to be borderline wars. The most notable challenges to the democratic peace include the American Civil War, the Spanish-American War, and the Boer War.

Developing Versus Consolidated Democracy

An important element of state proclivity to engage in international war is the stability of the involved state's domestic political environment. The difference between a stable and consolidated democracy and an unstable political institutional environment arguably has a significant impact on the warring behavior of the state in question.

Recall the three primary criteria used to identify and categorize democracies: executive recruitment, executive restraints, and political participation. The ideal democracy possesses the highest levels of all three categories, while the most extreme autocracy the lowest. A third distinct category, *anocracy*, is a state occupying the somewhat nebulous middle area between autocracy and democracy. The process of transitioning from any one of these regime types to the other can be termed either *democratization* or *autocratization*. States are considered to be

democratizing if, during a given period of time, they change from autocracy to either anocracy or democracy, or if they change from anocracy to democracy . . . autocratizing if they change from democracy to anocracy or autocracy, or from anocracy to autocracy. (Mansfield & Snyder, 1995, p. 9)

These processes have a profound impact on the internal stability of states and, in turn, their warring nature.

Democracies are significantly different from their autocratic counterparts, making for difficult transitions. Autocracies are characterized by the centralization of decision-making processes and very often the forceful silencing of opposition. As such, when states move from autocracy to democracy, there are suddenly myriad localized and sometimes national political groups that had previously been politically excluded. Local and national elites are compelled to compete for their respective groups' interests. Very often, this competition can either overwhelm the newly formed democratic institutions, leading to domestic turmoil, or elites can engage in risky foreign policy in attempts to bolster their domestic support (e.g., Mansfield & Snyder, 1995). Both domestic turmoil and foreign policy maneuvering from the democratization process arguably increase the potential for war.

Although seemingly intuitive, statistically speaking, the argument that regime change increases war proclivity is inconclusive. Mansfield and Snyder (1995) argue that states undergoing democratization show a roughly 60% increased probability of being involved in war relative to stable counterparts. However, scholars have had difficulty replicating their assessment. Neither Maoz and Abdolali (1989), Enterline (1998), nor Thompson and Tucker (1997) find any evidence that democratization increases either the possibility of international war or even lower level international dispute. Rather, these authors independently argue that only autocratization increases the likelihood of interstate dispute. Given the ultimate rarity with which democracies fight (recall that there are no incidences of democracies engaging in war with one another and minimal cases of war with autocracies), there is room to debate the absolute impact of democratization on violence.

Firing the First Round: Who Initiates War

A glaring question remains to be answered. Within mixed dyads composed of both a democracy and an autocracy, is one state more prone to initiating war than the other? By addressing this question, scholars can directly assess the theory of democratic pacifism in terms of war initiation and examine if democracies are, in general, more peaceful than are other types of states or if they are dragged into wars by other violent states.

On further investigation of this phenomenon, democracies instigate a significantly smaller proportion of wars with autocracies than do autocracies with democracies (Reiter & Stam, 2003). "Democracies are not significantly more likely to target dictatorships than vice versa, but dictatorships are significantly more likely to target democracies" (p. 336). To further extrapolate this point, between 1816 and 1990, democracies initiated a total of 15 wars, whereas autocratic regimes (dictatorships and oligarchs) initiated 42 wars. However, it is worth noting that not only do democracies initiate a sizable number of wars with

autocracies (Bueno de Mesquita & Lalman, 1992), but also they tend to win those that they instigate (Lake, 1992; Reiter & Stam, 2002). Democracies win 93% of the wars they initiate, whereas dictatorships win roughly 60% of conflicts they initiate (Reiter & Stam, 2002).

Democratic and autocratic victory records, it is argued, are a byproduct of the respective democratic and autocratic systems. Democratic leaders, knowing that they can be easily removed from office, choose to involve themselves in short, relatively bloodless wars that can be won fairly easily. Thus, there is a strategic selection process within democracies where wars that can be won are instigated, and those that would potentially drag on are overlooked. Contrarily, autocratic leaders have less fear of popular reprisal for their decisions to engage in war. Thus, autocracies instigate more wars than do democracies, and they tend to lose a fair number more of those wars they instigate given the limited consequences of losing war and a lessened need for strategic choosing of winnable wars.

One is thus led to the conclusion that there is a significant amount of war being fought in the international system at any given time and that democracies play their fair share. Democracies instigate and win a sizable number of wars against autocracies, and although the overall number of wars involving democracies is substantial, democracies simply do not fight each other.

International Context: Democracy and the Cold War

As discussed thus far, claims of the absence of international war between democracies are dependent on definitions of war and democracy, the stage of democratic development, and recognition that democracies are very willing to engage autocracies in open warfare. This chapter must now address the fact that democratic war is also dependent on (a) the number of democracies within the international system and (b) the willingness of democracies to fight one another under the auspices of a unified, nondemocratic threat.

At the end of World War II in 1945, there were as few as 19 recognized democracies in the world, many of which were less than a decade old (Marshall & Jaggers, 2007). As of 2009, the number of democracies had grown to approximately 120, and they now represent the majority of the states in the international system. However one chooses to perceive this trend, an indisputable fact is that democratic proliferation is a relatively recent phenomenon, and until only incredibly recently there were very few democracies in the world.

Critical to this point is the timing of this spike in democratic states. From 1945 until 1991, the international system was bifurcated between two economic and ideologically competing super powers, the United States and the Soviet Union. The era of the cold war pitted capitalist democracies on the side of the United States and

centrally organized communist states with the Soviet Union. During this period, the United States and the Soviet Union enforced peace between their respective ideological groups. Further, the United Nations was largely incapacitated during the cold war by the presence of U.S. and Soviet veto power in the UN Security Council and pressure on their respective blocs to vote uniformly. As such, with such a minimal number of democracies in the world and the relatively recent adoption of democracy as a system of state governance, democracies have had relatively little opportunity to engage one another in war. Coupled with the advent of the cold war and United States pressure to avoid conflict with other democracies, one could make the plausible argument that the democratic peace is merely an artifact of (a) the low level of contact between an equally low level of democracies and (b) cold war power politics.

The Realist Counterargument to the Democratic Peace

Perhaps the dominant theoretical approach to international relations in the latter half of the 20th century was realism (see Chapter 37, "Realism and Neorealism," within this volume for a detailed discussion). For realists, the necessity of state survival in the international system precludes domestic politics, and regardless of whether a regime is democratic or autocratic, a state will act similarly to other states. Among others, Waltz (1979) famously declared states to be "billiard balls" in that given identical international circumstances, states will respond identically.

The reality is, however, that states do not respond identically to similar scenarios. Realism presents no reasoned argument why democracies should not go to war with the same frequency as all other states. As billiard balls, the international system should drive them to conflict with the same propensity as all other states, and yet it does not. Democracies alone do not engage in war, and they do so only when paired together.

Policy Implications of Democratic Peacefulness

Rarely does a scholarly observation gain such momentum within the policy community. U.S. presidents have routinely stated that international peace is dependent on the presence of democracy. Among others, Bill Clinton famously remarked in his 1994 State of the Union Address, "Ultimately, the best strategy to ensure our security and to build a durable peace is to support the advance of democracy elsewhere. Democracies don't attack each other" (cited in Reiter & Stam, 2002, p. 1). More recent rhetoric by U.S. leaders pertaining to the spread of democracy in the Middle East further weights this point.

Can Democracy Be Sold?

Given the propensity for Western political leaders to draw attention to the importance of democracy on the world stage, it is prudent to address the question of whether democracy can be successfully supported or imposed in previously nondemocratic nations. This question is particularly important given the proclivity for the United States and international organizations such as the International Monetary Fund and United Nations to explicitly pursue the systematic democratization of states. The question must be asked: Can democracy be sold?

Enterline and Greig (2005) argue that democracy, imposed by an outside power, can have grave implications for localized peace. Their argument is contingent on the “brightness” of the imposed democracy. A bright democracy is analogous to a highly consolidated democratic state, while a dim democracy is equivalent to an anocracy and may be quite politically unstable. Results are telling. Bright democracies reduce localized interstate war and are particularly effective at reducing war waged by states that share political borders with the new democracy. However, bright democracies have no democratizing impact on the region. Contrary to bright democracies, dim democracies actually increase localized instability and reduce the likelihood that regional states will democratize. In the instance a state shares a border with a dim imposed democracy, they have an increased likelihood of becoming involved in war and lessened chances of democratizing.

Given these results, politicians would be wise to consider the potential repercussions of imposing democratic institutions in states or regions where democracy has not traditionally existed. Where imposed democracies “burn brightly [and] . . . reflect strong democratic institutions” (Enterline & Greig, 2005, pp. 1089–1090) localized violence may decrease. The potential for a dim democratic imposition is a very glaring reality, however, and may actually increase localized violence and slow the spread of democracy: the exact opposite reaction politicians may desire.

Future Directions

Future scholarship on the democratic peace would do well to explore several areas of research. First, there has been a glaring absence of work investigating democratic processes and warring behavior. Work on developing or transitional democracy provides only a rough proxy for the limited ability of weak democracies to function properly. This should not inhibit scholars from exploring different manners in which democracies process and institutionalize means of conflict resolution. For example, majoritarian democracies such as the United Kingdom approach policy construction in a manner quite distinct from a consensus democracy, which is exercised, for example, in Switzerland.

Second, the willingness of democracies to fight must be addressed. Gartzke (1998) argues that if democracies vote in a similar fashion in the United Nations, they should have nothing to fight about given the vote’s reflection on state preferences. Although the evidence is intriguing that democracies vote in a very similar fashion, this argument is hindered by the aforementioned cold-war context argument. It is thus not surprising that democracies have voted in a consistent manner, thereby reflecting unified preferences and avoiding conflict. Continued scholarship on convergence of democratic preferences in the post-cold war era would be highly informative, and studies outside of the United Nations would inform the current understanding.

Third, scholars must address the notion that democracies are a relatively new phenomenon in the international system and that a continued increase in the number of democracies could have substantial consequences. Over time, the democratic fraternity has increased in number, and democracies are increasingly likely to have direct contact with one another. To assume that the democratic peace is a law when much of the time period under consideration was largely devoid of democracy and to allow the argument to persevere in the face of an ever-increasing number of democracies would be a tremendous oversight. Just as an increase in a state’s population gives rise to an increased probability that any one person could partake in an extreme act of violence, it is likely that as the number of democracies in the international system increases, so does the proclivity for interaction and deviant cases of war between democracies. Further, as a derivative of the spread of democracy, scholars must also be cognizant that the ever increasing number of democracies could, in fact, be driving increases in instability between the decreasing number and thus increasingly isolated autocratic nations (Ray, 1995).

Conclusion

Although democracy’s international advancement has been relatively recent, the observation that democracies rarely, if ever, fight has attracted the attention of academic and policy minds alike. Many studies, empirical and purely theoretical, have added substantial weight to the argument. Yet one must not assume this research agenda to be complete. Many questions remain in this area of inquiry, focusing not only on the statistical rarity or definition of democracy but also on democracy’s constitution and how its spread (imposed and not) will impact the world stage in the future. Rigorous examination of international context, democratic process, and democratic interest are needed.

Inconsistencies within the democratic peace argument, such as the presence of international cold-war alliances (e.g., NATO) and the concurrent explosion of democratic states post-World War II, or the willingness of democracies to attack nondemocracies with regularity, also require

further examination. Although scholars disagree as to the exact causes behind the democratic peace, only continued examination will tell if the democratic peace is truly worthy to bear the heavy yoke of law or if it is merely an aberration.

References and Further Readings

- Barbieri, K. (1996). Economic interdependence: A path to peace or a source of interstate conflict? *Journal of Peace Research*, 33(1), 29–49.
- Beck, N. (2001). Time series cross sectional data: What have we learned in the past few years? *Annual Review of Political Science*, 4, 271–293.
- Bennett, S. D. (2006). Toward a continuous specification of the democracy autocracy connection. *International Studies Quarterly*, 50, 313–338.
- Bremer, S. A. (1992). Dangerous dyads: Conditions affecting the likelihood of interstate war, 1816–1965. *Journal of Conflict Resolution*, 36, 309–341.
- Bueno de Mesquita, B., & Lalman D. (1992). *War and reason*. New Haven, CT: Yale University Press.
- Cohen, C. (1971). *Democracy*. New York: Free Press.
- Crescenzi, M. (2003). Economic exit, interdependence, and conflict. *Journal of Politics*, 65, 809–832.
- Dahl, R. (1998). *On democracy*. New Haven, CT: Yale University Press.
- Dixon, W. J. (1994). Democracy and the peaceful settlement of international conflict. *American Political Science Review*, 88, 14–32.
- Doyle, M. W. (1983). Kant, liberal legacies, and foreign affairs. *Philosophy and Public Affairs*, 12, 205–235.
- Enterline, A. J. (1998). Regime changes and interstate conflict, 1816–1992. *Political Research Quarterly*, 51, 385–409.
- Enterline, A. J., & Greig, M. J. (2005). Beacons of hope? The impact of imposed democracy on regional peace, democracy, and prosperity. *American Journal of Political Science*, 67, 1075–1099.
- Fearon, J. D. (1994). Domestic political audiences and the escalation of international disputes. *American Political Science Review*, 88, 577–592.
- Gartzke, E. (1998). Kant we all just get along? Opportunity, will ingness, and the origins of the democratic peace. *American Journal of Political Science*, 42, 1–27.
- Gartzke, E. (2007). The capitalist peace. *American Journal of Political Science*, 51, 166–191.
- Gates, S., Hegre, H., Jones, M. P., & Strand, H. (2006). Institutional inconsistency and political instability: Polity duration, 1800–2000. *American Journal of Political Science*, 50, 893–908.
- Goemans, H. E. (2000). *War and punishment*. Princeton, NJ: Princeton University Press.
- Kant, I. (1903). *Perpetual peace: A philosophical essay*. London: Swan Sonnenschein. (Original work published 1795)
- Lake, D. A. (1992). Powerful pacifists: Democratic states and war. *American Political Science Review*, 86, 24–37.
- Lake, D. A. (2003). The new sovereignty in international relations. *International Studies Review*, 5, 303–323.
- Lake, D. A. (2008). The state and international relations. In C. Reus Smit & D. Snidal (Eds.), *Oxford handbook of international relations* (pp. 41–61). New York: Oxford University Press.
- Levy, J. S. (1988). Domestic politics and war. *Journal of Interdisciplinary History*, 18, 653–673.
- Mansfield, E., & Snyder, J. (1995). Democratization and the danger of war. *International Security*, 20, 5–38.
- Maoz, Z., & Abdolali, N. (1989). Regime types and international conflict, 1816–1976. *Journal of Conflict Resolution*, 33, 3–35.
- Marshall, M. G., & Jaggers, K. (2007). *Polity IV country reports 2007*. Retrieved December 10, 2009, from <http://www.systemicpeace.org/polity/polity06.htm>
- Mitchell, S. M., & Prins, B. C. (1999). Beyond territorial contiguity: Issues at stake in democratic militarized interstate disputes. *International Studies Quarterly*, 43, 169–183.
- Ray, J. L. (1995). *Democracy and international conflict: An evaluation of the democratic peace proposition*. Columbia: University of South Carolina Press.
- Reiter, D., & Stam, A. C. (2002). *Democracies at war*. Princeton, NJ: Princeton University Press.
- Reiter, D., & Stam, A. C. (2003). Identifying the culprit: Democracy, dictatorship, and dispute initiation. *American Political Science Review*, 97, 333–337.
- Russett, J., & Oneal, J. (2001). *Triangulating peace: Democracy, interdependence, and international organizations*. New York: W. W. Norton.
- Sarkees, M. R. (2000). The correlates of war data on war: An update to 1997. *Conflict Management and Peace Science*, 18, 123–144.
- Schumpeter, J. A. (1955). *Imperialism and social classes*. New York: Meridian.
- Thompson, W. R., & Tucker, R. (1997). A tale of two democratic peace critiques. *Journal of Conflict Resolution*, 41, 428–454.
- Waltz, K. (1979). *Theory of international politics*. Boston: McGraw Hill.

GLOBAL POLITICS OF RESOURCES AND RENTIERISM

GIRAY SADIK

Gazi University

Natural resources of countries are expected to contribute to their economic and political development. Natural resources can be regarded as assets for economic development. Then the economic development can constitute a base for political development. In the case of industrialized Western democracies such as the countries of the European Union and the United States, economic and political development have been taking place hand in hand. Accordingly, the expectation for the industrializing countries has been developed along these lines: The ones with greater natural resources are expected to develop faster.

Yet, especially in the Middle East, where almost all Arab countries have substantial oil wealth, resistance to democratization has remained strong. Thus, the expectation of natural resource abundance to become an asset for economic and political development has not been fulfilled at all. This anomaly can have both cultural and economic explanations. Although a cultural explanation for this democracy gap refers to historical legacies of Arabs, since they have never been democratic, and in part the role of Islam in Arab society, an economic explanation emphasizes oil wealth as the main barrier to democracy. Although one needs to keep in mind that most nations had nondemocratic times in their histories, the rentier state theory aims to explain that lack of political development in resource-rich states in modern times. This explanation is based

on the assumption that not only the resource availability but also states' methods to extract their resources (i.e., tax revenue versus oil contracts and foreign aid) have an impact on the political development of states. According to the rentier state theory, "Authoritarianism prevails where profits from natural resource exports displace taxes in government revenues" (Luciani, 1990a, p. 77). Since the ruler is no longer in need of taxes to maintain its authority, then he (male in all Arab autocracies) has no economic incentive to democratize his country. Thus, rentierism has a profoundly negative effect on the prospects of democratization.

Although the underlying assumption of rentier state theory looks simple, its theoretical reformulations and applications need to be further explained in order to understand how rentierism works in some cases and fails to work the same way in some others. To this end, this chapter starts by explaining the basic logic of the rentier state theory, then discusses the theory in light of the empirical findings, and finally outlines the policy implications and further research venues with respect to the rentier state theory.

Theory

Before elaborating on the empirical applications and political implications of rentierism, it is essential to understand its theoretical foundation. The core claim of the rentier

state theory is that rentierism is bad for democracy. The explanation for this foundational claim is mainly economic, or more specifically from political economy perspective. To be more specific, it is not the resource abundance that hinders political development but rather the rentier behavior of the rulers toward their citizens, which the governments can afford only in the presence of external rents derived from oil, minerals, or foreign aid. Thus, rentier state theory is at the crossroads of where rentierism as an economic structure of the country's economy hinders its political development and democratization.

According to the rentier state theory, the economic explanation for the democracy gap emphasizes oil wealth as the main obstacle to democracy. For Hazem Beblawi, himself of Middle Eastern origin, where most oil-rich states display rentier characteristics, and one of the pioneer thinkers of the rentier state theory (with Giacomo Luciani), resource-rich states like the oil countries in the Middle East are often termed *rentier states*, because the following is true:

They derive most of their revenues from external rents. Rents are paid by foreign actors, accrue directly to the state, and only a small fraction of the population is engaged in the generation of this rent, while the rest nevertheless might benefit from the distribution and use of it. (Beblawi, 1990, p. 23)

Now, once the logic of the rentier state theory is clarified, one needs to know the answer to these questions: What does the rentier state mean? What are the attributes of a rentier state? What is necessary to be considered a rentier state by the rentier state theory? Although they may seem to be typical "What is the yardstick" type questions, it is important to address them in the very beginning of the theoretical explanation, so as to understand to what types of states the rentier state theory refers. What are the common characteristics of the rentier states on which the rentier state theory is based? To what extent can the rentier state theory capture and explain those characteristics?

According to Luciani (1990a), "A rentier state is a country whose government typically receives at least 40% of its revenues in the form of rent" (p. 72). Luciani's clarification has been a source of wide consensus among the political scientists studying the rentier state theory and its political repercussions. In addition, Hazem Beblawi's definition clearly explains the core idea of what makes a state a rentier. In a rentier state, for Beblawi (1990), the following is true:

1. The rents come from abroad,
2. The rents accrue to the government directly, and
3. Only a few are engaged in the generation of this rent (wealth), the majority being only involved in the distribution or utilization of it. (p. 87)

It is important to note that it is especially the third point that makes the use of the term *rent* appropriate. Hence, this

concept (i.e., rentierism) in some ways resembles the usage of the term (i.e., rent) in classical political economy: Rents are not generated by productive human activity but instead by the scarcity value of natural endowments. In this case, clearly the global scarcity of oil presented an unprecedented income potential for those oil-rich countries, so they have been able to survive on these oil riches with almost no productive activity. Rentier state theory goes a step ahead of this economic explanation and claims that oil hinders not only economic development but also political development. According to Beblawi (1990), this hindrance of oil riches extends even beyond the oil-rich Arab states in the Middle East. Therefore, for Beblawi, "The oil phenomenon has cut across the whole of the Arab world, oil rich and oil poor. Arab oil states have played a major role in propagating a new pattern of behavior, i.e., rentier behavior" (p. 91).

To better explain the rentier state theory, it is important to delineate the causal mechanisms of what Beblawi (1990) calls rentier behavior. What are the sources of the rentier behavior? How has the rentier behavior been implemented and sustained throughout the Arab Middle East? What are the sociopolitical interactions between the rentierism and Arab society? How has the rentier behavior affected the societal dynamics and democratization in the Arab Middle East?

To answer all of the aforementioned questions, one needs to understand the causal mechanism of the rentier behavior from the perspective of the rentier state theory scholars. For example, according to Ross (2001), the causal mechanisms underlying the theory are of three sorts:

1. How does the state collect revenue?
2. How does the state spend revenues?
3. How does the rentier wealth distort social structure, and prevent changes that promote democracy? (p. 325)

Beginning with the very first one, which can hence be regarded as the root of the causal mechanism: How does the state collect its revenue? The very traditional answer to this question in political economy is taxation. From the rentier state theory view, though, oil riches of the state release the governments of the oil-rich Arab states from depending on the taxes from their populations to gain their revenues. Therefore, the scholars of the rentier state theory claim that the absence of taxation "releases the state from the accountability ordinarily exacted by domestic appropriation of surplus . . . the state may be virtually completely autonomous from its society, winning popular acquiescence through distribution rather than support through taxation and representation" (Anderson, 1987, p. 10). A number of rentier state theory scholars point out that this argument draws heavily on the lessons of European history, in which it is widely thought that taxation had a major role in the development of democratic

institutions (Ross, 2001). In addition, according to Luciani (1990a), “The lack of taxation undercuts the organization of citizens based on economic interests, and makes religious and cultural organization paramount” (p. 89).

After explaining how the very source (i.e., oil rents) of the government revenue hinders the development of democratic society and institutions, it is also important to consider how the distribution of this wealth is likely to affect the prospects for political development and democratization. Thus, the second causal mechanism aims to explain this question: How does the state spend its revenues? Put simply, from rentier state theory perspective, rentierism increases the capacity of the state to both buy off, and to repress, the opposition. As a result, according to Wiktowicz (1999), the combined effect of these two mechanisms is often thought to produce a “rentier social contract” in which “the state provides goods and services to society . . . while society provides state officials with a degree of autonomy in decision-making” (p. 608).

Although the first two sets of causal mechanisms are principally related with how the government collects and distributes the oil rents and their political implications, the third mechanism concentrates more on the societal influences of the rentier behavior. A third set of causal explanations in the literature holds that rentier wealth distorts social structure by preventing changes that promote democracy as compared with countries that follow a more standard development trajectory. Thus, while the first two causal explanations are state centered, the third focuses on how rents affect society in general and thus is more about natural resource dependence than rentierism specifically.

To better understand the last causal mechanism, one needs to be aware of the difference between the natural resource dependence and rentierism. Although there is a close correlation between rentierism and natural resource dependence, they are not the same thing. To delineate the difference between the two, one needs to consider how they are measured so as to know specifically to what exactly each refers. The standard measure of natural resource dependence is natural resource exports as a percentage of GDP: This assessment is classically used in the debates of the economic consequences of exporting natural resources. By contrast, rentierism focuses on rents in government revenues and is the preferred measure in the rentier state theory. For example, in most rentier states, rents from oil and other resources (e.g., natural minerals or foreign aid) constitute more than the half of their governments’ budgets. In practice, dependence on oil exports is observed almost always along with rentierism. Furthermore, the effects of other resources with rentier potential are examined at the end of this section.

After understanding the core logic and the causal mechanisms of the rentier state theory, one needs to know about the foundational studies and their key contributions to the development of the rentier state theory. The publication of the edited book *The Arab State* by Giacomo Luciani (1990b)

marked the beginning of renewed and intensified discussions about the first strand in the literature of the rentier state theory concentrating on the political economy approach to the study of the Arab state. The most important and influential contributions to this book were the articles by Giacomo Luciani (1990a), titled “Allocation Versus Production States,” and Hazem Beblawi (1990), titled “The Rentier State in the Arab World,” in which the authors argued that those states that derived a substantial part of their revenues from the outside world and whose functioning of the political system depends to a large degree on accruing external revenues that can be classified as rents demonstrated a remarkable different political dynamic than other (i.e., productive) states. Rents were defined as “the income derived from the gift of nature” (Luciani, 1990a, p. 38) and are thus usually understood to be income accrued from the export of natural resources, especially oil and gas. In addition, Beblawi and Luciani argued that the rentier effects are not confined to the oil-exporting states alone. This is first due to the fact that to a limited but still significant extent the rents of the oil state have been recycled to the non-oil Arab states through migrant workers’ remittances, through transit fees, and through aid. Second, the authors stressed that external rents may also be conceived of as bilateral or multilateral foreign-aid payments, such as foreign development assistance or military assistance, which are termed *strategic rents* (Luciani). Furthermore, other theoretical considerations in this regard include the contribution by Luciani (1995) to the edited volume titled *Political Liberalization and Democratization in the Arab World, Volume 1: Theoretical Considerations*.

The aforementioned two theoretical contributions by Luciani (1990a) and Beblawi (1990) have become benchmarks in the literature of the rentier state theory, and their political economy approach has been used to explain the lack of political development in the Arab Middle East. Their political economy approach has also been applied as the basis for single-country studies, cross-country studies, and thematic studies, which are discussed in detail in the next section on the applications and the empirical evidence from the rentier state theory.

Applications and Empirical Evidence

The core claim of the rentier state theory is that rentierism hinders the development of democratic society and political institutions. This theoretical argument has been advanced in a number of case studies and theoretic pieces by Lisa Anderson (1987) and Giacomo Luciani (1990a). As the previous details of which discussed, the theoretical contributions by Luciani and Beblawi (1990) soon became benchmarks in the literature, and their political economy approach was used as the basis for single-country studies, cross-country studies, and thematic studies. Thematically,

the 1990s saw the emergence of a vast literature that analyzed the issue of economic liberalization and privatization in the countries of the Arab Middle East from a political economy perspective. These studies include cross-country analysis, such as the edited volumes by Henri Barkey (1992) as well as several single-country studies. Furthermore, others have tried to make a synthesis of these single-country studies by coming up with some general observations in the field of the rentier state theory. In their cross-country study on the Middle East, they come to the conclusion that the degree and kind of rentierism will determine the level of economic liberalization of the state.

Applications of the aforementioned theoretical considerations to single-country cases include Rex Brynen's (1992) work on Jordan and several contributions in the edited volume *Democracy Without Democrats* by Ghassam Salamé, most notably the chapters by Abdelbaki Hermassi (1994) on the Maghreb and Volker Perthes (1994) on Syria. In addition, a host of other country case studies were published in 1998 in the second volume of *Political Liberalization and Democratization in the Arab World, Volume 2: Empirical Considerations* (Brynen, Korany, & Noble, 1998). All of these empirical studies confirmed the theoretical claim made by Luciani that the rentier nature of the state is a strong factor in discouraging democratization in states that have access to a significant oil rent. This rentier nature of the state can still be a factor in the some Middle Eastern countries with limited resources such as Egypt, Syria, and Jordan. Strategic rents (i.e., foreign aid, especially to Syria and Egypt by the USSR during the cold war) and the remittances from their workers earned from richer Arab countries such as Saudi Arabia and United Arab Emirates have fostered the development of the rentier state in the absence of rich oil reserves in these countries. More recently, a time-series cross-national study using data from 113 states between 1971 and 1997 confirmed these initial empirical studies and showed that oil exports are strongly associated with authoritarian rule (Ross, 2001). In the same vein, the political economy approach with its focus on the rentier effect has resulted in academic studies that center on the political dynamics at large in certain individual countries of the Middle East. Furthermore, several authors (Gause, 1994; Moss, Pettersson, & Walle, 2006) have taken a more thematically oriented focus and have stressed the effects of rentierism on a state's foreign policies, on a state's human rights policy, or on aspects of political succession in authoritarian states.

In addition to the aforementioned thematic country studies, Ross (2001), in one of the major published large-number, cross-regional tests of the rentier state theory, finds that "the oil-impedes-democracy is both valid and statistically robust . . . oil does hurt democracy" (p. 356). Barro also finds that oil exporters are less likely to be democratic. As a result, the rentier state theory steps to a more general level, advancing beyond being the theory used for the individual country-case explanations for the

lack of democratization in the Arab Middle East. Consequently, these applications of the rentier state theory and the empirical evidence gathered as a result of these studies are likely to be helpful in unveiling the political implications of rentierism in the Middle East and beyond.

Foreign Aid as Rent?

In search of the explanatory potential of the rentier state theory beyond the Middle Eastern oil-rich states, one needs to explore its applicability to the states with limited resources. In other words, it is necessary to inquire whether other resources such as foreign aid might function as a source of rents for the state as well. This would not only help explain to what degree the concept of the rentier state is limited to oil-rich states but also help to test whether the rentier state theory is more generally applicable beyond the oil-rich states of the Middle East that already have various common characteristics ranging from religion to political culture.

Addressing the aforementioned questions is important when assessing the validity of the rentier state theory beyond the oil-rich states of the Middle East. From the rentierism perspective,

The core proposition is that there was a set of strong synergies between (a) the degree of dependence of rulers on tax revenue, (b) the emergence of representative government, and (c) the strength and resilience of the state in the context of interstate competition, especially war. (Moore, 2004, p. 297)

Especially with respect to the first two propositions (a and b), for the rentier logic to hold, any resource other than the tax revenue extracted by states as a result of production should have the similar rentier effects. A growing literature argues the following:

A range of deficiencies and pathologies in the political constitution of many states in the "South" can be traced to a high level of dependence on *natural resource rents* (especially oil and minerals) and *strategic rents* (especially foreign aid), rather than taxes. (Moore 2004, p. 297)

In financial terms,

The dominant type of strategic rent in the contemporary world is the many forms of development aid. Development aid has in recent decades been increasingly concentrated on the poorest countries, and has always, for geostrategic reasons, been given more generously to small countries. (Moore, 2004, p. 302)

For example, since foreign aid has become the major source income for many sub-Saharan countries, its negative effects on state institutions have been observed in recent studies. From the rentier point of view, the impact of foreign aid dependence on the relationship between state and citizen is especially worth considering. In their

joint report Moss et al. (2006) observed that “states which can raise a substantial proportion of their revenues from the international community are less accountable to their citizens and under less pressure to maintain popular legitimacy” (p. 1). Thus, foreign aid has become a source of rent for many small countries in sub-Saharan Africa. Their rulers in turn have become reluctant to provide the incentives to cultivate effective public institutions, since the rulers themselves were lacking that economic pressure (i.e., the need for tax revenue). As a result, “substantial increases in aid inflows over a sustained period could have a harmful effect on institutional development in sub-Saharan Africa” (p. 3).

The core premise of the rentier state theory is based on the assumption that “it matters whether a state relies on taxes from extractive industries, agricultural production, foreign aid, remittances, or international borrowing because these different sources of revenues have powerful (and quite different) impact on the state’s institutional development” (Karl, 1997, p. 34). From the rentier state theory perspective, at the core of this different impact is whether states rely on their own extractive capacity (i.e., tax revenue) or some sort of external financing such as revenue shares from foreign oil companies or foreign aid. Although foreign aid and oil revenue can have different effects on the rentier nature of a state, in the end because the state lacks the capacity to generate its own production (i.e., tax revenue from private businesses), it is unlikely to have that democratic pressure akin to the one most European states have experienced in the aftermath of the Industrial Revolution. In this regard, foreign aid does have a theoretically (i.e., from a rentier theory perspective) similar effect as a spoiler for the governments, enabling them to afford ruling their societies without the need for taxes and thus providing leverage to those rentier governments to turn a blind eye to public demands for representation and democratic rights.

Policy Implications

The rentier state theory is the most prominent theoretic paradigm in the study of the comparative politics of the Middle Eastern Arab states, and it has increasingly been applied to the study of natural resource exporters in other regions of the world. Therefore, the rentier state theory and its applications have important policy implications in the study of the domestic and foreign policies of the Arab states in the Middle East.

Many scholars of the Middle East and the third world in general have argued that the availability of external rents has led to the development of what Hartmut Elsenhans has termed a *state class* or for what William Reno has coined the term *shadow state*. In essence, both terms describe the same phenomena—namely, a self-serving ruling elite that has control over the vast natural resources of the country that provide the elite a financial basis for the government

revenue, instead of the tax revenue collected by consolidated democracies from their citizens. As a result, by controlling the major source of revenue (i.e., oil and gas), this small elite gained and maintained its political power. For that reason, the core argument of the rentier state theory starts from within the state so one can observe how this oil wealth is generated and distributed domestically in order to maintain the rule of the small elite by buying off or suppressing any democratic alternatives to their rule.

In the contemporary Arab states in the Middle East, one finds nation states, which on the one hand are composed of heavy bureaucratic state institutions but on the other hand hold only a weak legitimacy with respect to their societies. The infrastructural power of these Arab states as well as their capacity to actively control political outcomes independent of societal constraints is limited. Hence, the small elite has created a set of state institutions and heavy bureaucracy to institutionalize its political power based on the continuing oil revenues, instead of taxing the domestically generated products. As a result, these rentier Arab states in the Middle East can be considered weak states. A weak state can be defined as a country characterized by weak state capacity, weak state legitimacy, or both, and thus fragile state institutions; hence, the term *fragile state* is also used to describe the same concept. Given the apparent fragility of the Arab territorial state on the one hand and the fact that these states are here to stay, Bahgat Korany, another scholar from the region, has clearly pointed out the contradictions of the Arab territorial state). As a consequence, most of the academic research on the policy implications of the rentier state theory in the Middle Eastern Arab states veers around the puzzle about the resilience and the persistence of the contemporary Arab states. Thus, the rentier state theory aims to explain the durability of the rentier Arab states, despite the fact that there is a lack of political development and democratization in the Arab Middle East.

Among the most significant policy implications of the rentier state theory is its explanation of the changes (or lack of change) in the regime types of the Middle Eastern Arab states, which are mostly absolutist monarchies with some reference to Islamic customs such as Sharia. The core of the argument in this respect is the fact that economic well-being is regarded as the most important requisite for democracy (Lipset, 1959). In a number of quantitative studies, per capita GDP (or GNP) has emerged as the most stable predictor of democracy (Lipset, Seong, & Torres, 1993). As a result, all of these studies demonstrated that prosperous nations are more likely to be governed democratically than poor ones. This association between economic well-being (measured as per capita GDP or GNP) has therefore been established as “one of the most powerful and robust relationships in the study of comparative national development” (Diamond, 1992, p. 125). In terms of their per capita income, the majority of the Arab states in the Middle East are rather wealthy, so that, according to the simple version of the previously outlined modernization

theory, these Arab states should already have turned toward more democratic forms of government. However, although there have recently been small steps of political opening in small countries like Bahrain, Qatar, and Oman, the absolute monarchy remains the dominant form of government among the Arab states in the region.

Rentier state theory's principal explanation of the previously outlined democracy gap despite high levels of per capita GDP is that the characteristics of a rentier state make democracy less probable. Thus, the rentier state theory has important policy implications at the very domestic level for these Arab states in the Middle East. First, because of the revenue derived from the sale of oil, the governments do not need to collect high taxes; in fact, they often don't collect taxes at all. As a consequence of that, the governments in those countries are confronted with fewer demands for accountability and representation by the public or can afford to ignore them (Gause, 1994; Ross, 2001). Since the population has had very limited to nonexistent contribution to the generation of the government revenue, the governments of the Arab rentier states have become virtually unbound in terms of public spending, distributing the oil wealth in a way that maintains their rule. In the history of the productive states, the rulers' attempts to raise taxes have often led to demands for political participation and accountability, hence the motto "No taxation without representation" seems to reflect the political reality in the European world. In contrast, in the resource-rich states of the Middle East, this motto seems to be inverted to read, "No representation, without taxation." In addition to the absence of widespread public demands of political accountability in the Arab Middle East, oil rents have also been instrumental to the rentier governments as a principal source for the state spending on patronage, subventions, and subsidies. For example, free education and health care are provided to the population. In general, people indulged like this without having to pay anything (and usually by receiving certain payments from the state on top) are satisfied with their lives and feel no need for political participation. Within this type of rentier system, there is virtually no incentive to form associations or interest groups, and some of the governments even take deliberate action to depoliticize the population, which has never been politically active in a European sense. Even if political groups averse to the government still happen to form, owing to the oil rents, the government is still able to prevent them from becoming too strong. In addition, oil rent enables the Arab governments to spend more on internal security and sustain a large coercive apparatus (Ross, 2001).

The rentier state theory argues that rentier states stand in contrast to states that have to rely on domestic resource extraction. Thus, the policy implications of the rentier state theory are in contrast to the theories seeking to explain the development among the European states. The rentier states display a particular path to state formation that by and large defies the European path of state formation: Natural resource dependence (mainly oil dependence) has created

weak states that are autonomous from societal demands and that do not rely on domestic taxation. As a result, in contrast to the European states, the state formation in the rentier states has not been accompanied by political accountability and transparency. In rentier states, the expenditure side of public revenues is most clearly linked to a state-building agenda of creating societal peace through political acquiescence, which aims to maintain the authority of the small ruling elite in the long term.

When analyzing the policy implications of rentierism, one needs to take into account resource-rich countries beyond the Middle East. Does one see the similar rentier effects in Nigeria, Venezuela, Algeria, Gabon, or Indonesia? Can one say that rentierism is limited only to oil-rich countries? Can other natural resource-rich countries demonstrate similar rentier characteristics? All in all, can rentier state theory explain significant policy implications of resource-rich countries beyond the Middle East?

In search of an answer to these questions, Ross (2001) used pooled time-series cross-national data from 113 states between 1971 and 1997. His findings demonstrated that "oil exports are strongly associated with authoritarian rule; that this effect is not limited to the Middle East; and that other types of mineral exports have a similar antidemocratic effect, while other types of commodity exports do not" (p. 14). Thus, the resource abundance not only generates tax revenues and royalties for governments but also, more critically, creates a dependence of governments' budgets on natural resource rents in the Middle East and beyond. For instance, "Indonesia's profits sharing contracts reserves up to 90% of oil profits for the government, Venezuela's Energy Information Administration takes 85–94% of oil profits and 60% of Mobil's earnings in Nigeria accrue to the Federal government" (Wantchekon, 2000, p. 72). None of these countries are in the Middle East, they are not even close to each other, but they do demonstrate similar rentier effects. How countries as unrelated as Saudi Arabia, Venezuela, Gabon, Iran, Nigeria, Algeria, and Indonesia ended up in profound economic and political crisis calls for an explanation. These countries, according to Karl (1999),

are heterogeneous in virtually every respect except oil: they are physically diverse (Algeria is more than 100 times larger than tiny Kuwait) and demographically different (Indonesia's population is 132 times that of Qatar); they vary in their oil reserves (Saudi Arabia has 265 times as much as Gabon). (p. 37)

Despite various differences ranging from geography and history to demography and political culture, these states have been at the juncture of underdevelopment and political crisis. In this respect, the rentier effect helps to explain how so many diverse and rich countries fail to realize political development. Bates (2000) neatly explains this rentier effect:

It is useful to contrast the conduct of governments in resource rich nations with that of governments in nations less favorably endowed. In both, governments search for revenues; but they

do so in different ways. Those in resource rich economies tend to secure revenues by extracting them; those in resource poor nations, by promoting the creation of wealth. Differences in natural endowments thus appear to shape the behavior of governments. (p. 4)

The fact that the rentier effect plays a critical role in explaining the government behavior in a diverse range of countries worldwide provides another significant reason for the applicability of the rentier state theory for policy analysis in different political and geographical contexts.

Future Directions

From the rentier state theory perspective, as long as resources are abundant, autocratic regimes that derive their revenue from oil rents are in a comfortable position, and they can postpone the democratization indefinitely. This theoretical perspective is derived from the very core argument of the rentier state theory that oil rents are not only sources of valuable government revenues but also, more significantly, essential sources of political power. For this reason, oil rents can be used not only as substitutes for the lack of economic productivity in the Arab Middle East but also to compensate for a lack of political development and institutions such as political representation and accountability.

Since these Arab states in the Middle East rely on the oil rents for their economic and political sustainability, the rentier state theory also relies on the oil rents and on ongoing rentier behavior to preserve the validity of its explanations. Thus, once the oil is depleted, it's rentier behavior, and thus, rentier behavior explanations are likely to be abandoned. Thus, the challenge for those oil-rich Arab states is yet to start once their oil is gone. As Yousef (2004) observes, the economic performance of the Middle Eastern rentier states was at its peak from the 1950s to the 1970s, while oil prices were consistently high, whereas from the 1980s onwards, declining oil prices and a more competitive international environment led to a decline in growth rates and public revenues. This observation is no surprise at all for the scholars of the rentier state theory; as a matter of fact, it can be regarded as an expected early warning call, since it is common knowledge that natural resources are not infinite. As result, even the small oil emirates that are still able to maintain a comparably high standard of living are suffering from problems of an economy based on natural resource abundance: underperformance in long-run GDP growth, rising unemployment, and the lack of foreign investment. States whose resources are limited are especially likely to soon face the need to compensate for lower rents. There are basically two possibilities to act: Governments may choose to adapt to lower income by cutting expenditure, raising taxes, or practicing deficit spending and thus try to maintain the status quo as long as possible. The alternative is to restructure and diversify the economy (Luciani, 1994, 1995).

Economic diversification is executed by building up new industries, trading companies, or a banking sector. But economic modernization will bring about social changes, too. Education levels, occupational specialization, and interaction with foreign economies will rise. Since the regime is no longer able to buy consensus by distributing goods, services, and incomes in exchange for little or nothing, it will need some kind of legitimization, and the citizens, in turn, will demand accountability and will want to influence political decisions that affect their lives and the business sector they work in. To increase its legitimacy, the regime might ponder complementing economic liberalization with political reforms in a democratic direction. All of these economy-led (i.e., depleting natural resources) sociopolitical challenges constitute new venues for future research for the scholars of the rentier state theory. In addition, in almost all studies in the literature of the rentier state theory, the emphasis is squarely on the negative effects of rents on democracy; the idea of balancing the positive and negative effects of rentierism on democratization has yet to be fully explored.

Conclusion

The rentier state theory uses the political economy explanation for the democracy gap in the Arab Middle East, which emphasizes oil wealth as the main barrier to democracy. In sum, from the rentier state theory perspective, oil rents make possible a fairly high standard of living for the people, without any productive economic activity. In turn, oil rents enable the governments to keep their publics politically demobilized, either by fiscal generosity or by repression, and do not bring about the social changes that usually lead to political mobilization in favor of democracy. As quantitative studies demonstrate, oil wealth has indeed a strong negative impact on the level of democracy (Ross, 2001).

The large and considerable amount of state revenues accruing to rentier states in the form of external oil rents gives the governments additional resources and thus serve to reduce the state's need to extract money from its society. As a result, weak states that are virtually independent from their respective societies emerged and are sustained as long as the oil rents remain. Considering the fact that oil reserves are limited, the prospects for the rentier states once the oil is gone remain to be among the areas for future research.

References and Further Readings

- Anderson, L. (1987). The state in the Middle East and North Africa. *Comparative Politics*, 20(1), 1–18.
- Anderson, L. (1990). Policy making and theory building: American political science and the Islamic Middle East. In H. Sharabi (Ed.), *Theory, politics and the Arab world: Critical responses* (pp. 52–80). London: Routledge.
- Ankie, H. (1997). *Globalization and the post colonial world*. London: Macmillan.

- Auty, R. M. (1993). *Sustaining development in mineral economies: The resource curse thesis*. London: Routledge.
- Barkey, H. J. (Ed.). (1992). *The politics of economic reform in the Middle East*. New York: St. Martin's.
- Bates, R. (2000). *Prosperity and violence: The political economy of development*. New York: W. W. Norton.
- Beblawi, H. (1990). The rentier state in the Arab world. In G. Luciani (Ed.), *The Arab state* (pp. 85-98). London: Routledge.
- Brynen, R. (1992). Economic crisis and post rentier democratization in the Arab world: The case of Jordan. *Canadian Journal of Political Science*, 25(1), 69-97.
- Brynen, R., Korany, B., & Noble, P. (Eds.). (1995). *Political liberalization and democratization in the Arab world, vol. 1: Theoretical perspectives*. Boulder, CO: Lynne Rienner.
- Brynen, R., Korany, B., & Noble, P. (Eds.). (1998). *Political liberalization and democratization in the Arab world, vol. 2: Comparative experiences*. Boulder, CO: Lynne Rienner.
- Chaudhry, K. A. (1990). Economic liberalization in oil exporting countries: Iraq and Saudi Arabia. In I. Harik & D. J. Sullivan (Eds.), *Privatization and liberalization in the Middle East* (pp. 145-66). Bloomington: Indiana University Press.
- Chaudhry, K. A. (1997). *The price of wealth: Economies and institutions in the Middle East*. Ithaca: Cornell University Press.
- Diamond, L. (1992). Economic development and democracy reconsidered. In G. Marks & L. Diamond (Eds.), *Reexamining democracy: Essays in honor of Seymour Martin Lipset* (pp. 93-139). Newbury Park, CA: Sage.
- Gause, G. (1994). *Oil monarchies: Domestic and security challenges in the Arab Gulf states*. New York: Council on Foreign Relations.
- Helliwell, J. F. (1994). Empirical linkages between democracy and economic growth. *British Journal of Political Science*, 24(2), 225-248.
- Hermassi, A. (1994). Socio economic change and political implications: The Maghreb. In G. Salamé (Ed.), *Democracy without democrats? The renewal of politics in the Muslim world* (pp. 200-226). London: Tauris.
- Karl, T. L. (1997). *The paradox of plenty: Oil booms and petro states*. Berkeley: University of California Press.
- Karl, T. L. (1999). The perils of the petro state: Reflections on the paradox of plenty. *Journal of International Affairs*, 53(1), 25-42.
- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53, 69-105.
- Lipset, S. M., Seong, K. R., & Torres, J. C. (1993). A comparative analysis of the social requisites of democracy. *International Social Science Journal*, 136, 155-175.
- Luciani, G. (1988). Economic foundations of democracy and authoritarianism: The Arab world in comparative perspective. *Arab Studies Quarterly*, 10(4), 457-475.
- Luciani, G. (1990a). Allocation vs. production states: A theoretical framework. In G. Luciani (Ed.), *The Arab state* (pp. 65-84). London: Routledge.
- Luciani, G. (Ed.). (1990b). *The Arab state*. London: Routledge.
- Luciani, G. (1994). The oil rent, the fiscal crisis of the state and democratization. In G. Salamé (Ed.), *Democracy without democrats? The renewal of politics in the Muslim world* (pp. 130-155). London: Tauris.
- Luciani, G. (1995). Resources, revenues, and authoritarianism in the Arab world: Beyond the rentier state? In R. Brynen, B. Korany, & P. Noble (Eds.), *Political liberalization and democratization in the Arab world, vol. 1: Theoretical perspectives* (pp. 211-227). Boulder, CO: Lynne Rienner.
- Mahdavy, H. (1970). The patterns and problems of economic development in rentier states: The case of Iran. In M. A. Cook (Ed.), *Studies in the economic history of the Middle East* (pp. 428-467). London: Oxford University Press.
- Mann, M. (1993). *The sources of social power, vol. II*. Cambridge, UK: Cambridge University Press.
- Moore, M. (2004). Revenues, state formation, and the quality of governance in developing countries. *International Political Science Review*, 25(3), 297-319.
- Moss, T., Pettersson, G., & Walle, N. (2006). *An aid institutions paradox? A review essay on aid dependency and state building in sub Saharan Africa* (Working Paper No. 74). Washington, DC: Center for Global Development.
- Okruhlik, G. (1999). Rentier wealth, unruly law, and the rise of opposition: The political economy of rentier states. *Comparative Politics*, 31(3), 295-315.
- Omeje, K. (Ed.). (2008). *Extractive economies and conflicts in the global south: Multi regional perspectives on rentier politics*. Aldershot, UK: Ashgate.
- Perthes, V. (1994). The private sector, economic liberalization, and the prospects of democratization: The case of Syria and some other Arab countries. In G. Salamé (Ed.), *Democracy without democrats? The renewal of politics in the Muslim world* (pp. 243-269). London: Tauris.
- Ricardo, D. (1921). *On the principles of political economy and taxation*. London: John Murray.
- Ross, M. L. (2001). Does oil hinder democracy? *World Politics*, 53(3), 325-361.
- Sandbakken, C. (2006). The limits to democracy posed by oil rentier states: The cases of Algeria, Nigeria and Libya. *Democratization*, 13(1), 135-152.
- Wantchekon, L. (2000). *Why do resource dependent countries have authoritarian governments?* New Haven, CT: Yale University Press.
- Waterbury, J. (1997). From social contract to extraction contracts: The political economy of authoritarianism and democracy. In J. P. Entelis (Ed.), *Islam, democracy, and the state in North Africa* (pp. 141-176). Bloomington: Indiana University Press.
- Wiktorowicz, Q. (1999). The limits of democracy in the Middle East: The case of Jordan. *Middle East Journal*, 53(4), 606-620.
- Yousef, T. M. (2004). Development, growth and policy reform in the Middle East and North Africa since 1950. *Journal of Economic Perspectives*, 18(3), 91-116.

48

COMPLEX INTERDEPENDENCE AND GLOBALIZATION

DJAMEL EDDINE LAOUISSET

Canada Consulting Group, Inc.

Complex interdependence and globalization have become core concepts in most academics' minds. Nonetheless, disagreement on working definitions has led to conceptions centering mostly on the ideas of economic, cultural, and technological interdependence and interconnectedness.

Most scholars acknowledge that the concept of globalization has the merit of amalgamating social organizations unto one global society. As a matter of fact, highly intricate relations of push and pull forces are producing simultaneous integration, degeneration and divergence, order and chaos at the interregional or transnational level.

These are unifying and conflict-ridden forces within globalization, which can generate remarkable opportunities for affluence, peace, and democracy but also threats for divergence, business supremacy, and lack of consideration for world citizens and civilizations.

Globalization: A Theory of Expectations

Globalization has always been equated by many analysts with economic interdependence (Bhagwati, 2007). Indeed, nowadays, the extent and level of global economic relations appears to be unparalleled in world history, mainly in terms of the immense quantity of capital flows.

Emergent countries, also, are progressively becoming a part of global business and investment flows. Modern-day

economic globalization models imply a new international division of labor reflecting the new global economy, whereby economic and financial integration do not remain solely concentrated amid the industrialized countries of North America, Europe, and East Asia (Amin, 1996). As a matter of fact, global capital has not stimulated enough policy homogenization, and significant differences in economic structuring subsist.

Multinational corporations, for instance, which are seen by many as globalization's leading agents, remain mostly active in their country of origin. The debate as to whether economic globalization will aggravate economic inequalities or contribute to advancing economic justice lead to a considerable amount of literature on the impact of globalization on wealth distribution amid both most-developed countries and *less developed countries* (Friedman & Kaplan, 2002).

What has been observed is that the effect of globalization has been both positive and negative, and the effect was largely dependent on a range of internal and external variables. Countries that are trade partners, for instance, because they actually trade more with each other, are less likely to enter into conflictual relations with one another.

In fact, greater ties from interdependence have been argued to lead to both greater cooperation and conflict. It has been observed that greater trade led to peace and peace leads to greater trade. Many nongovernmental organizations, and global civil society in general, are resisting some aspects of globalization, advocating that human

rights and environmental protection should be aligned with economic efficiency.

Restructuring the global system or just reforming the existing one has always been a fundamental question. Antiglobalization activists seek a more equitable distribution of wealth, greater societal participation, and immediate solutions for the global environmental crisis (Fukuyama, 2001).

Global Networks: Applications and Empirical Evidence

While interdependence refers to a state of affairs, globalization implies that something is ever increasing. Globalization can be contrasted with localization, nationalization, or regionalization; it refers to the contraction of distance on a large scale.

As for deglobalization, it refers to the decline of globalism. As such, globalism refers to networks of connections, involving not only regional networks but networks of interdependence at multicontinental distances. Interdependence refers therefore to any situation characterized by dialectical effects among countries or actors in different countries (Oneal & Russett, 1997).

Scholars Keohane and Nye (1998) have extended the analysis of transnational relations away from classical political economy to include contentious international politics. These scholars initiated a new way of looking at the world of politics via observing the relationships between economics and politics, and the patterns of institutionalized international cooperation while retaining key realist insights about the roles that power and interests play in world politics.

The multiple linkages take the form of flows of capital and goods, as well as information and ideas between the center and the periphery. Interdependence and globalism are both multidimensional phenomena, essentially defined in economic terms, as if economic factors are the sole determinant of globalism. In fact, economic globalism involves long-distance flows of goods, services, and capital, as well as the information, perceptions, and organizational processes carried by market exchange. However, social and cultural globalism involves the movement of ideas, information, images, and peoples behind it, such as the movement of religions or the diffusion of scientific knowledge.

A central feature of social globalism involves the imitation of one society's practices and institutions by others, what some sociologists refer to as *isomorphism*, *demonstration effect*, or *transfer of social technology* (Laouisset, 2009), and in doing so flow across geographical and political boundaries, hence transforming societies and markets and affecting the consciousness of individuals, their personal identities, and their attitudes toward culture and politics.

Social globalism is also impacted by the phenomenon of network effects. For professional economists, the term

network effects refers to situations where a product becomes more valuable once many people use it, such as the Internet. As such, a knowledge-based economy generates powerful spillover effects, which spread rapidly, trigger additional innovation, and lead to chain reactions of new inventions (Stieglitz, 2006). Moreover, as interdependence and globalism become thicker, systemic relationships among different networks intensifies.

The worldwide impact of the financial crisis that began in the United States in 2008 illustrates the extent of these network interconnections. Unexpectedly, what first appeared as an isolated real estate-related crisis had severe global effects. It generated losses everywhere, and for instance, in the case of the United Arab Emirates, particularly the Emirate of Dubai, it prompted emergency meetings at the highest level of local finance and huge rescue packages orchestrated by the Emirate of Abu Dhabi, and it led eventually to a general loss of confidence in the United Arab Emirates real estate market and the efficacy of the country's economic model (Davidson, 2009).

The financial crisis of 2008 and 2009 was not the first to be global in magnitude. The Wall Street crisis in 1929 triggered a worldwide financial crisis and depression. But some characteristics of the 2008 and 2009 crisis differentiate it from preceding ones. Most economists, governments, and international financial institutions failed to predict the crisis, and intricate new financial instruments made it difficult to understand (Eichengreen, 2006). Even countries such as Canada that had been praised for their sound economic policies and performance were no less vulnerable to the financial contamination triggered by speculative offensives and capricious changes in market feeling. Sheer magnitude, complexity, and speed distinguish contemporary globalization from earlier periods (Greenspan, 2007).

The ever-increasing thickness of globalism and the density of networks of interdependence is not just a variation in degree, but thickness means that diverse relationships of interdependence interconnect more intensely at more points. Hence, the effects of events in one geographical area, on one dimension, can have profound effects in other geographical areas, on other dimensions.

As in chaos theory, small events taking place in a specific time and space can have catastrophic effects in another time and space. Such systems being difficult to comprehend, their effects are therefore difficult to predict. As a result, globalism will likely be accompanied by omnipresent ambiguity.

There will be constant competition between increased complexity and uncertainty and efforts by governments, market participants, and other actors to comprehend and manage these increasingly complex interconnected systems (Burtless, 2007).

Globalization, therefore, does not merely have an effect on governance; it is in turn as much affected by governance, and scholars prefer to use the term *global governance*, in contrast to the traditional meaning of the term

governance, to denote the regulation of interdependent relations in the absence of an overarching political authority obviously benefiting from an extensive use of networks (Rosenau, 2007).

Global Threats and Challenges: Policy Implications

International trade, international capital flows, and transnational corporations have a direct impact on our lives, and never before have there been so many poor and so many disparities (Kothari, 1997). This fact alone constitutes a legitimate basis for revolts and conflicts. Critics of globalization denounce lowering of real wages, deregulation, and lowering social benefits, as well as privatizations of public services, relocating factories, and loosening the grip of organized labor.

When directed at the Structural Adjustment Programs of the 1980s and 1990s, these charges are seemingly justified. Governments were basically losing their economic authority, even within their own countries. Poor countries, living in considerable monetary instability, cannot stop poverty from growing, nor can they pull their economies out of dependency.

Globalization appears to bring greater insecurity for labor, making unemployment and underemployment chronic social ills and leading to greater inequality in income levels. The South has its North: the rich elite who is part of the world's top income bracket. Similarly, the North has its South: the immigrants and the unemployed (Rodrik, 2002).

National institutions in MDCs have not yet developed appropriate policies, measures, and mechanisms that truly protect people from the ravages of free market capitalism. The costs of this system are so visible in terms of social and racial fragmentation, criminality, and the collapse of family and community.

One of the most negative impacts of globalization is environmental degradation, a very serious international threat to the extent that environmental and health-related concerns challenge the narrowed dimension of the state-centered model, hence redefine the concept of united community, promote the idea of global public goods, and address global public ills via collaborative networks (Mittelman, 2002). However, when we address globalization and globalism, we perceive globalism as a rooted historical trend and globalization as the process of this ever-increasing globalism (Scholte, 1997).

The question to be asked is not how old globalism is, but rather how thin or thick it is in space and time. Thin globalization can provide an economic and cultural connection between trading partners, and traded goods impact only a small number of consumers. In contrast, thick globalization (Held, 2007) engages large and uninterrupted long-distance flows, affecting the lives of many consumers. For instance, the 2008 and 2009 operations of global financial

markets had an ill effect on everyone, mainly because globalization is the process by which globalism become increasingly thick.

The degree of thickening of globalism is giving rise to three changes in degree and in kind, such as increased density of networks, increased institutional velocity, and increased transnational participation (Greenspan, 2007). This participation and consequent transparency has been a byproduct of the information revolution that is at the center of economic and social globalization; it has made possible the transnational organization of work and the growth of markets, thereby facilitating a new international division of labor.

The division of labor being limited by the extent of the market (Smith, 1776), the information revolution had a major impact on attempts to expand globalism. However, globalization divides and polarizes fragments.

While complex interdependence between national economies and societies has increased significantly, social relations between societies have increased considerably. Globalization, the term used to summarize the ever more complex interdependence among nations and societies in terms of financial flows, trade, industry, and communication, has likewise lead to some economic growth and contributed noticeably to a reduction in poverty in the Asian and Pacific region. However, the impact of these aspects of globalization has not been uniform, and socioeconomic disparities between and within some countries and areas in Asia and the Pacific have sometimes widened during that same period (Owens, Baylis, & Smith, 2008).

To ease these harmful aspects of globalization, a greater consideration to social development is recommended, both as the critical goal of economic development and as a means of achieving such development. For economic development to persist, social development must be self-sustaining, and social and economic policies must be integrated for either to be effective and sustaining. For this to happen, an adequate amount of investment in human resources is required, such as the provision of education, health, shelter, and sanitation. Productive employment and economic empowerment are the most valuable means for citizens to contribute to social and economic development.

Citizens must play a part also in the decision-making processes that shape their lives at the local and national levels (Laouisset, 2009). Global citizens are also interested in the way conflicts are resolved, how international politics are managed by states, and obviously, this interest exposes them to both situations of realism and idealism as paradigms in international affairs. Realism is a perception named as such because authors of this theory believed it more realistic.

By rejecting idealism (Kant, 2003), these people believe peace is more possible by following the path they recommend and hence do not believe cooperation is achievable in the manner idealists wish for (Locke, 1997). Being the core components of the system, states interact in a seemingly anarchic system, since they lack central authority

such as an idealized world government (Angell, 1912). It is therefore difficult to build cooperative structures, and it makes conflict inevitable in world affairs. However, humans do have the capacity for rationality, and that capacity is what can allow statesmen to construct a system that can peacefully deal with conflicts of interest.

Rational states wish to maximize relative power, and this is achieved whenever systemic status quo is maintained. Rational leaders, for instance U.S. government leaders, recognized that even though human rights were not respected in Algeria during the 1990s political crisis, support for the Algerian military dictatorial regime was necessary. Hence, realists can sacrifice democracy and international morality for foreign policy gains and perceived national interest and security, hiding sometimes behind foreign sovereignty principles. Realists argue that in the current international system one must accept that other states will have their own way of doing things (Kennan, 1984).

This necessitates accepting that a state may be dictatorial; as long as they are not trying to alter the system, you can deal with them (Kissinger, 1973). The theory of complex interdependence has argued that realism is becoming archaic, liberalism is a superior approach to international affairs. This argument has modified conventional liberalism by emphasizing international institutions as facilitated global free trade and cooperation, and this approach has been labeled neo-liberalism (Keohane, 1998).

Realists need to address the impact of economics and especially economic interdependence on power relations. Whenever two countries are interdependent, then the conservative approach to look at competition as a zero-sum game is sometimes seen as old-fashioned.

But real life is not a zero-sum game, our best prospects are usually in cooperative efforts, and it does in fact move from zero sum to positive sum. Hence, the game theoretic approach has clarified the conditions required for the evolution and persistence of cooperation and shown how Darwinian natural selection can lead to complex behavior, including notions of morality, fairness, and justice. It is shown that the nature of self-interest is more profound than previously considered, and that behavior that seems altruistic may, in a broader view, be individually beneficial (Axelrod, 1997).

Moreover, war and military conflict would be even more costly since they would not only destroy people and things but would also break the entire economic systems of both or all states involved. This makes military power less important than the past and economic links a stronger variable.

Future Conceptual Directions

The concept of complex interdependence is being replaced by the concept of globalization, since the latter reflects

both interdependence and its consequent ramifications. The concept of interdependence has been developed on the premises that inequalities and injustices appear as a result of history and social behavior (Huntington, 1996).

The establishment of global citizenship allows each citizen of each nation to be fully aware of his or her rights and duties not only in the face of national government and law but also in the face of international law and order. The concept of interdependent globalization seems much too complex, but globalization provides exciting opportunities for those committed to human solidarity and justice as well as immense new challenges for globalization in solidarity and globalization without marginalization. Globalization is a process, not a static event, and it involves complex, interdependent networks (Krugman & Venables, 1995).

The concept of globalization belongs to the 1990s social, economic, and political literature, just as interdependence belongs to the 1970s scene (Amin, 1996). Like all concepts meant to cover complex phenomena, both interdependence and globalization have various meanings. As governance structures are established at the global level in order to deal effectively with the increasing number of global issues, conflicts have also emerged as how to make international organizations more democratically accountable.

To deal with such developments, states have found it useful to erect international organizations and endow them with significant decision-making authority. It became increasingly obvious that real authority has been hence transferred to international organizations and other nonstate actors. Civil society pressure groups have had, so far, a major effect on nation-states and international organizations (Hirst & Grahame, 2001). Given the close relationship between globalization and technological innovation, the literature has also examined how new technologies will affect concepts of citizenship and democracy in allowing citizens to challenge authoritarian governments and truly participate in advanced industrial democracies.

The substitution of blue-collar labor intensive technologies with white-collar highly capital-intensive technologies, the communication revolution, and the proliferation of global media may also make it easy to customize the information citizens receive, thereby lessening opposing views and increasing social discipline and political silence. However, the emergence of nongovernmental organizations and global social movements as new political actors provide evidence for a global civil society and a new culture.

This cultural globalization phenomenon, seen also as a Westernization process, is for the most part driven by corporations, rather than countries, and aims at the broadening of consumer culture (Hirst & Grahame, 2001). However, the greatest challenge of globalization lies in the cultural sphere. In fact, there is the corruption of local cultures to create one global culture of consumerism via television and the Internet and the feeling of alienation and confusion

it produces in some sectors because of the breakdown of the local cultural fabric (Appadurai, 1990).

We all have our own values, cultures, and social life to root and ground us, and indeed in a time of painful transition, culture is the most important jewel to safeguard. However, globalization has also generated impressive solidarity in transnational social movements that work for genuinely implementing the United Nations Charter, the Universal Declaration of Human Rights, protection of the environment, the protection of refugees, and the establishment of an International Criminal Tribunal.

State sovereignty is being redefined by the forces of globalization and international cooperation whereby states are now widely understood to be instruments at the service of their people, and not the contrary (Bohman, 1999). Hence, individuals are first of all members of the world-wide community, with rights that derive from their humanity, and second, they are members of the communities of existing nation-states, calling hence for the globalization of citizenship.

Conclusion

Globalization is an attempt to describe new features of the interaction of people and institutions in our postmodern world. Among the central challenges of effective leadership in the 21st century are understanding and responding to the patterns of interdependence that now increasingly characterize the social, economic, and political spheres. It is no longer possible to address effectively local concerns except in complex, global contexts. Neither is it possible to respond aptly to confounding, large-scale global issues without assessing and attending to local subtexts.

The scale and depth of interdependence characteristic of the contemporary world are bringing a variety of social, economic, and political communities and institutions into unprecedented close relationships. But it is also bringing about the interfusion of widely differing cultures and interests. The choices confronting contemporary societies thus cannot be limited to deciding on factual solutions under the assumption of essentially shared values but necessarily entail negotiating broad assent on both common norms and the meaning of beneficial change. "We cannot expect to solve today's problems with the same thinking that created them" (Einstein, 1954).

The search for solutions to the global problems needs to be local, regional, and global so as to unite the world so that its abundant skills and resources may be directed toward the challenges of the real problems of humankind. It is difficult to deny that significant changes are occurring throughout the world.

Globalization is seen in the export of economics, politics, and culture. It seems reasonable to assume that this new sharing of information and culture would help people

in the world to understand each other better, which would encourage peace and help nations to live and work more successfully together.

By sharing these aspects of our lives, we are creating a global community. It seems possible that one day there will be a global culture and one global economy. The future effects of globalization are and will still continue to be widely debated. Will globalization end economic, political, and cultural differences throughout the world? Or will it serve only to widen and exacerbate these differences? We believe that globalization does have the potential to end longstanding national differences, but we believe it is happening at a slower pace than it may appear on the surface.

We also believe that for globalization to be ultimately successful in this endeavor and maximize winners, there are still many challenges to come that must be overcome, among them the necessary sensitivity and understanding of other cultures, traditions, and belief systems.

Culture should not be sacrificed at the altar of globalization. If handled with care, globalization could one day greatly reduce differences in political, economic, and cultural systems, while preserving what is unique and special in various countries (Laouisset, 2009).

If well monitored, globalization can lead to global prosperity and a bigger market of goods through a wider freedom of movement for goods and people in the global economy.

References and Further Readings

- Amin, S. (1996). The challenge of globalization. *Review of International Political Economy*, 3(2), 216-259.
- Angell, N. (1912). *The great illusion*. London: Knickerbocker Press.
- Appadurai, A. (1990). Disjuncture and difference in the global culture economy. *Theory, Culture, and Society*, 7, 295-310.
- Axelrod, R. (1997). *The complexity of cooperation*. Princeton, NJ: Princeton University Press.
- Bhagwati, J. (2007). *In defense of globalization*. Oxford, UK: Oxford University Press.
- Bohman, J. (1999). International regimes and democratic governance. *International Affairs*, 75(3), 499-513.
- Burtless, G. (2007). *Globalization and income polarization in developed countries*. Washington, DC: Brookings Institution.
- Davidson, C. (2009). *Abu Dhabi: Oil and beyond*. New York: Columbia University Press.
- Eichengreen, B. (2006). *Global imbalances*. Washington, DC: Brookings Institution.
- Einstein, A. (1954). *Ideas and opinions*. New York: Random House.
- Friedman, T., & Kaplan, R. (2002). States of discord. *Foreign Policy*, 129, 64-70.
- Fukuyama, F. (2001). Their target. *Newsweek*, 138(25), 42-48.
- Greenspan, A. (2007). *The age of turbulence*. New York: Penguin.
- Held, D., & McGrew, A. (2007). *Globalization & anti globalization: Beyond the great divide*. Cambridge, UK: Polity.
- Hirst, P., & Grahame, T. (2001). *Globalization in question*. Malden, MA: Polity Press.

- Huntington, S. (1996). *The clash of civilizations*. New York: Touchstone.
- Kant, I. (2003). *Perpetual peace*. London: Humphrey.
- Kennan, G. (1984). *American diplomacy*. Chicago: University of Chicago Press.
- Keohane, K. O., & Nye, J. S. (1998, September/October). Power and interdependence in the information age. *Foreign Affairs*, 81 94.
- Kissinger, H. (1973). *A world restored*. Gloucester, MA: Smith.
- Kothari, R. (1997). Globalization. *Alternatives*, 22, 227 267.
- Krugman, P., & Venables, A. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics*, 110(4), 857 880.
- Laouisset, D. (2009). *Managerial technology transfer*. New York: Nova.
- Locke, J. (1997). *An essay concerning human understanding*. New York: Woolhouse.
- March, J., & Olsen, J. (1999). The institutional dynamics of international political orders. In P. Katzenstein, R. Keohane, & S. Krasner (Eds.), *Exploration and contestation in the study of world politics* (pp. 303 330). Cambridge: MIT Press.
- Mittelman, J. (2002). Globalization: An ascendant paradigm? *International Studies Perspectives*, 3(1), 1 14.
- Oneal, J., & Russett, B. M. (1997). The classical liberals were right. *International Studies Quarterly*, 41(2), 267 293.
- Owens, P., Baylis, J., & Smith, S. (2008). *The globalization of world politics*. Oxford, UK: Oxford University Press.
- Rodrik, D. (2002). *Feasible globalizations* (NBER Working Paper No. 9129). Cambridge, MA: National Bureau of Economic Research.
- Rosenau, J. (2007). *People count: The networked individual in world politics*. New York: Paradigm.
- Scholte, J. A. (1997). Global capitalism and the state. *International Affairs*, 73(3), 427 452.
- Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: Strahan & Cadell.
- Stieglitz, J. (2006). *Making globalization work*. New York: Penguin.

49

INTERNATIONAL POLITICAL ECONOMY AND TRADE

STEFAN FRITSCH

Bowling Green State University

International trade has become one of the most important issues in domestic as well as international politics in recent decades. Although a growing number of historically oriented studies (Abu-Lughod, 1989) have shown that trade has been a salient issue among empires, states, and cities for centuries, it has become such a critical contemporary issue because countries' economies are now, more than ever, open to trade flows. They thereby create complex interdependence, defined as mutual dependence, between national economies. Technological progress has resulted in dramatically falling transportation and communication costs, whereas various liberalization policies have freed the exchange of goods and services from various tariff and nontariff barriers.

Representing one major area of economic globalization, trade remains a controversial topic, as recent World Trade Organization (WTO) conferences and street demonstrations in Seattle and other cities have shown (Rosenau, 2007). The controversy surrounding trade stems from the fact that interest groups and the broader public view their welfare as being directly affected by trade policy. Although export-oriented companies and societal groups that profit from export exert pressure for global and regional liberalization agreements, domestically oriented firms and civil society groups oppose efforts to further liberalize trade and expand the authority of the WTO and regional trade agreements.

The goal of this chapter is to present the major theoretical discussions revolving around international trade, as well as to provide empirical evidence and highlight recent developments in the study of the political economy of international trade.

Theoretical Perspectives on International Trade

Liberalism

Liberal theorists of international political economy (IPE) generally view trade as a positive-sum game that provides mutual benefits to individuals, companies, and states. Although liberal trade theory has evolved considerably since the 18th century, the core assumptions as formulated by Adam Smith and David Ricardo still represent a major part of theoretical justifications for free trade. Smith (1776/1993) argued that gains from free trade result from absolute advantages:

If a foreign country can supply us with a commodity cheaper than we ourselves can make it, better buy it of them with some part of the produce of our own industry, employed in a way in which we have some advantage. (p. 573)

Ricardo's (1817/2006) theory builds on the theory of comparative advantages. It observes that free trade can be

beneficial, even in the absence of absolute advantages, if countries reallocate labor to sectors in which they have comparative advantages and then trade with others who also specialize in their respective areas. Shortly later, Mill (1848/2004) highlighted the fact that free trade primarily is beneficial not because of the revenue generated by exports but because of the cost savings experienced through the import of cheaper foreign products. Liberals therefore argue that specialization and trade benefit countries, even if one country has an absolute advantage in producing all the products traded.

Although the arguments of early liberal political economists proved to be influential, they built on the assumption that comparative advantages rest solely on differences in labor productivity. Comparative advantages also result from other production factors such as capital or natural resources. The Heckscher–Ohlin theory states that a country's comparative advantage depends on its relative abundance or scarcity of labor and capital. It has comparative advantages in products that make intensive use of the abundant factors while products using scarce factors will be less competitive (Ohlin, 1933/1967). Therefore, industrialized countries specialize in capital-intensive goods while less developed countries (LDCs) specialize in the production and export of labor-intensive goods.

Building on the Heckscher–Ohlin theory, Stolper and Samuelson (1994) developed the Stolper–Samuelson theory, which explains why domestic socioeconomic groups support or reject free trade. Free trade benefits abundantly endowed production factors and hurts poorly endowed factors. Therefore, owners of abundant production factors favor free trade while owners of scarcely endowed factors oppose it. Rogowski (1989) has shown that increasing or decreasing exposure to trade either intensifies class conflict or urban–rural conflict depending on the factor endowment of individual countries.

While Heckscher–Ohlin and Stolper–Samuelson focus on the factor endowment, sectoral or firm-based theories of trade preferences follow the Ricardo–Viner model, also called specific-factors model. The argument behind this model is that because at least one production factor is immobile, all factors tied to import-competing sectors potentially lose from free trade while those in export-oriented sectors win. The factor specificity, which refers to the question of how tied certain production factors are to specific sectors, is the key difference between the Stolper–Samuelson and Ricardo–Viner models (Alt, Frieden, Gilligan, Rodrik, & Rogowski, 1996).

Empirical studies have tested both models, either individually or combined. Although Irwin (1996) has found evidence in support of the Ricardo–Viner model, Scheve and Slaughter (1998) have investigated evidence in support of the Stolper–Samuelson type factor model. Several scholars have provided additional insights regarding the relationship between industry structure and preference for free trade or protectionism. While low-skill and labor-intensive industries, which face import penetration, are

usually associated with high protection, export-oriented industries and multinational corporations (MNCs) favor free trade (Milner, 1988).

Although the various liberal trade theories presented here have considerably influenced policy discussions of the last decades, they have also been criticized. For example, Leontieff (1953) found that the United States was highly successful in exporting labor-intensive goods during the 1950s, even though it was the most capital-rich country. Although liberal theories discussed here explain interindustry trade, they are not able to explain intraindustry and intrafirm trade because they assume products to be homogenous, whereas differentiated products increasingly are traded within the same industry group. Liberals reacted by developing theories that intraindustry trade provides benefits such as economies of scale, the satisfaction of varied consumer tastes, and the production of sophisticated manufactured products.

Economic Nationalism, Realism, and Neomercantilism

Mercantilism or economic nationalism was the dominating preindustrial economic policy and trade theory before the emergence of liberalism during the late 18th century. Between the 15th century and mid-18th century, mercantilism contributed significantly to the establishment of the modern state system through its emphasis on national power. Despite liberalism's relative dominance in the academic and public discussions, economic nationalism still remains influential today.

Similar to liberal economic theories, neomercantilism is an umbrella term for various strands of thought revolving around issues of trade and (state) power. Economic nationalists perceive trade as one among several instruments to increase a country's power position in the international system. While liberals see power and wealth as opposing goals, neomercantilists emphasize their complementary character. The equal consideration of power and wealth overcomes the economic reductionism of most liberal trade concepts and helps to refocus attention on the central role of states in the global political economy (Ashley, 1983).

States can use the revenue generated by mercantilist trade policy to finance armies or influence enemies and allies. Protectionist trade strategies, mainly tariff and nontariff barriers, have been the preferred instruments to limit foreign imports and maximize the export of domestically produced goods. Because trade never is perfectly symmetrical—that is, not all countries can have a positive balance of trade—trade relations will ultimately be characterized by power struggles and conflicts between states (Heckscher, 1934).

For mercantilists such as Hamilton (1791/1966), international trade based on country-specific comparative advantages results in reduced economic self-sufficiency and national security. To promote the United States' economic development, he recommended an emphasis on

industry over agriculture, economic self-sufficiency, government intervention, and protectionism. In his opinion, “Not only the wealth; but the independence and security of a Country, appear to be materially connected with the prosperity of manufactures” (Hamilton, p. 291). List (1916), a German representative of mercantilist trade theory, argued that “a nation which exchanges agricultural products for foreign manufactured goods is an individual with one arm, which is supported by a foreign arm” (p. 130). List pointed to the fact that even countries like Great Britain had switched to free trade strategies in the second half of the 19th century, only after they had achieved industrial and technological supremacy through protection of their infant industries. Once a country had caught up with more advanced nations through “artificial measure,” free trade would become the natural mode of operation. Thus, List was opposed to protectionist trade policies once a country had successfully industrialized. Realist thought considerably influenced states’ economic policies during the interwar period. To protect their national interests, states adopted protectionism, currency devaluation, and foreign exchange controls.

The economic depression of the interwar period, as well as the outbreak of World War II, provided the impetus for political leaders to fundamentally transform the world economic system after 1945. However, although the postwar international economic system represented by the Bretton Woods institutions (WTO, IMF, World Bank, and GATT) was based on liberal thought, economic nationalists or mercantilists continued to modify their theoretical concepts in order to adapt to the major developments in the international political economy of trade since 1945.

The major contribution of neomercantilist scholarship for the IPE of trade during the early 1980s has been the theory of hegemonic stability. This theory asserts that a global economic system is most likely to remain open and stable if a hegemonic state is willing and able to provide the necessary resources and leadership to convince other states that its policies are beneficial (Gilpin, 1987). According to most scholars, hegemonic conditions have occurred in only a few cases, including Portugal, Spain and the Netherlands until the end of the 18th century, Great Britain during the 19th century, and finally under U.S. leadership after World War II. Despite its considerable contributions to the field of IPE, hegemonic stability theory has also experienced various criticisms.

First, writers concerned with hegemony define the term in state-centric terms as a situation in which one powerful state controls or dominates the lesser states in the system and imposes its goals and rules in various policy areas. However, it does not tell how and what type of control hegemonic rule requires (Wallerstein, 1984).

Second, scholars have differing views regarding the character and goals of hegemonic leadership. While benevolent hegemons pursue the promotion of generalized benefits known as public goods through rewards, the mixed form aims at the realization of generalized and personal

benefits by employing positive as well as negative coercive methods. The exploitative hegemony serves pure self-interest of the hegemonic state and relies heavily on coercion. While liberals focus on the benevolent hegemon who is willing to maintain open and stable economic relations and thereby provides public goods from which nobody can be excluded, realists often portray hegemons as following their national self-interest (Grieco, 1988).

Finally, liberal institutionalists question the realist assumption that hegemony is necessary to maintain open trade relations. Liberal critiques argue that the international economic system can remain stable and open even though the hegemon, who initially supplied the regime, has declined. Instead, if demand among states for a specific regime is large enough, the incentives to collectively maintain an international liberal economic regime might remain large (Keohane, 1984).

Neomercantilists have also argued against liberal assumptions concerning (naturally given) comparative advantage. As proponents of strategic trade theory suggest, states can actually create comparative advantages through proactive intervention in the economy through industrial targeting (Krugman, 1986). Investments in prospective sectors combined with interventionist trade policies in the form of selective protectionism and liberalization help to develop new infant industries and to create competitive advantage up to the point where open competition with other countries seems possible. Contrary to liberal ideas about a reduced role of government in economic issues, economic nationalists emphasize the central role that state governments can play in the governance of external economic relations and the catch-up processes of late industrializers (Gerschenkron, 1962).

The so-called developmental state actively intervenes in domestic markets and external trade relations to generate competitive advantages for its firms in various sectors. As a result, those countries successfully move from the status of being LDCs to that of newly industrializing countries (NICs) or industrialized developed countries (DCs). The concept of the development state has also served realist and neomercantilist scholars to attack the assumption, held by many globalization scholars, that growing economic interdependence and transnationalization of national economics has resulted in an erosion of state authority in global economics. Instead, they argue that states will remain at the center of economic governance, domestically as well as globally (Hirst & Thompson, 1999).

Historical Structuralism

Marxists emphasize the importance of class relations for the international economic and political order. Class relations—capitalists or the bourgeoisie on the one side and the working class on the other—are basically conflictual. According to Marx and Engels (1948), “One fact is common to all past ages, viz, the exploitation of one part of society by the other” (p. 29). Under capitalism, private

owners of the means of production (capitalists) extract surplus value from wage laborers, who can offer only their labor power to earn a living. The surplus is converted into capital and reinvested into new means of production. However, the exploitation of labor and the necessity for capital accumulation combined with a steadily increasing portion of capital in the production process—that is, investments in new production technologies—lowers the rate of profit, since the only source for surplus value—labor—is diminished by technological progress. According to Marx, these developments would ultimately lead to overproduction and underconsumption, since fewer workers compete for jobs that barely earn income at the subsistence level, which causes purchasing power and consumption rates to decline (Marx, 1867/1990).

Eventually, the exploitation and repeating economic crisis would provide the necessary conditions for a revolution of the impoverished working class. The result would be a society in which the means of production would be owned and controlled collectively. Marx himself did not write systematically on international economic relations. His works, however, provided the basis for various theories, which all represent a historical-materialist perspective in their analysis of international trade.

For Lenin (1939), imperialism represented the highest stage of capitalism and explained the temporary survival of capitalism because colonies provided the imperial metropolises with an outlet for their overproduction, as well as sources of raw materials and agricultural products. Hilferding (1910/1981) observed that capitalists instrumentalize the state to impose high tariffs and other trade barriers in order to minimize imports and maximize exports. Contrary to liberal ideas, a strong state is not confined to a watchman function but intervenes in the economy and is the means to expand territory and acquire new colonies.

However, contrary to Lenin's assumptions, imperialism did not represent the final stage of capitalism marked by violent interstate competition for new territories and a delayed working-class revolution. Moreover, the positive effects of imperialism foreseen by Marx and Lenin—namely, the development of colonies through import of technology and capital—did not occur. Even after gaining independence, former colonies continued to depend on foreign capital and technology and continued to produce mainly raw materials and agricultural products. This led to major rifts within the Marxist approach to IPE (Biersteker, 1993).

Dependency theory gained considerable prominence during the 1960s. Dependency scholars assume that industrialized capitalist countries either neglect LDCs or prevent them from achieving economic development and autonomy (Frank, 1967). They reject Marx's or Lenin's view that developed countries serve less developed ones in the long run by exporting capitalism and instead argue that capitalist development fundamentally depends on the exploitation of LDCs.

Dependency theorists question liberal assumptions that everybody benefits from free trade and point to the

negative effects of declining terms of trade for LDCs. Terms of trade describe the ratio of the value of imports to the value of exports. Most countries exporting primary products and importing manufactured products experience negative terms of trade—that is, they have to export more and more products to purchase the same quantity of manufactured goods. However, the terms of trade for LDCs could change in the 21st century with continuously rising world population and growing scarcity of primary products.

LDCs' dependence on the export of primary products puts them at a continuous disadvantage since the demand for manufactured products increases with higher incomes, while the demand and prices for primary goods remain relatively constant. Therefore, Prebisch (1962) argued that LDCs should adopt import substitution industrialization (ISI) strategies, which involved tariff barriers and emphasis on domestic production of manufactured goods to replace foreign imports and protect domestic infant industries. More radical scholars called for severing of trade relations with developed countries.

However, two developments challenged dependency theory. First, ISI strategies in many Latin American countries clearly failed (Adler, 1986). Second, the successful economic development of several Southeast Asian countries since the 1960s provided empirical evidence that peripheral countries could industrialize by pursuing strategies based on foreign direct investment (FDI), the import of foreign technology and exports. Dependency studies reacted by introducing the concept termed *dependent development*—that is, under certain conditions LDCs can industrialize successfully by serving the interests of capitalist core countries, for example, by exporting less technologically sophisticated goods (Gereffi, 1983).

Moreover, the theory has been criticized for focusing almost exclusively on factors on the international system level, while neglecting domestic causes that contribute to underdevelopment, as well as for granting LDCs little autonomy in view of external challenges. Another criticism relates to an overemphasis on the relations of exchange over relations of production. Finally, dependency theorists have been accused of bias toward Western capitalism, while neglecting other forms of exploitation, for example, in the Eastern bloc (Clark & Bahry, 1983). Therefore, the popularity of dependency theory as an analytical framework for the relationship between trade and (under)development has declined remarkably since the mid-1980s, although authors concerned with economic development still continue to draw on specific aspects of this theory.

Neo-Gramscian theorists have primarily extended the concept of hegemony into the realm of culture and ideas, such as capitalism, free market and free trade ideology, market discipline, or American culture. Building on the concept of cultural or ideological hegemony originally developed by Gramsci, they analyze the establishment of a national and, in a later stage, transnational hegemonic bloc. A hegemonic bloc is composed of political elites, a transnational managerial class, and parts of the working

class. It is able to establish a bourgeois hegemony by gaining the active consent of subordinate classes based on shared values, ideologies, and material interests by providing socioeconomic benefits and supporting the establishment of labor unions. One example for a transnational historic bloc would be post-World War II U.S. hegemony, which was able to forge increasingly global support (especially after 1990) for institutions like the IMF, the World Bank, or the General Agreement on Tariffs and Trade (later the WTO), which together enshrine liberal norms and values revolving around free market economy, free trade, and liberalized financial markets (Cox, 1983).

Major Developments in the Global Political Economy of Trade

The Growth of Global Trade

For most of the postwar period, world trade has grown faster than world output, especially since the 1990s. World exports, measured as a proportion of world output, have tripled between 1950 and 1998. In 2003, this ratio stood at 29% and reached 27% in 2005, compared to 17% in 1990 and 12.5% in 1970. In absolute terms, world merchandise trade exceeded 10 trillion U.S. dollars in 2005, almost 65 times the value of world trade in 1963. Trade in services stood at 2.42 trillion U.S. dollars, almost a sevenfold increase from 1980 (WTO, 2006).

Contemporary trade involves more countries and sectors than ever before. For example, the number of participating countries in WTO negotiations grew from 23 in 1947 to 149 in 1999. Developing countries represent a growing share of world export markets, especially but not exclusively in manufactured products, which increased from 19.2% in 1970 to 32.1% in 2005 (United Nations Conference on Trade and Development [UNCTAD], 2005). During the postwar period, the ratio of exports in proportion to GDP for all countries increased from 5.5% in 1950 to 17.2% in 1998, and especially for many Organisation for Economic Co-operation and Development (OECD) and developing countries, it more than doubled (Kaplinsky, 2006). As Held and McGrew (2007) conclude, "Trade now reaches deeper into more sectors of national economies as an expanded array of goods and services have become tradeable" (p. 75).

The Impact of Regionalism on Global Trade

Scholars interested in the geographical patterns of global trade have found that the global political economy of trade is marked by several diverging developments. As described previously, trade has become a global phenomenon in that it involves more countries and practically every world region. However, trade is not evenly distributed. It remains highly concentrated in the OECD countries, which account for 65% of the world merchandise trade and a

small number of East Asian countries that represent most of the developing countries' trade (WTO, 2006). Yet the developed countries' dominance has become diluted since the 1990s, mainly by the emergence of new trading powers like China, India, and Brazil, causing a new global division of labor and intensified competition through trade (UNCTAD, 2005).

This new global division of labor has been caused by (a) massive shifts of manufacturing capacities to the NICs in East Asia, mainly by MNCs' FDI activities; (b) falling transportation and communication costs; and (c) the liberalization of trade and related FDI (UNCTAD, 1996). As a result, the export of manufactured products by developing countries has doubled from 31.4% in 1980 to 68.1% in 2005. Trade between developing countries (south-south trade) has almost doubled from 22.9% to 40.9% of their total exports (UNCTAD, 2005) but remains highly concentrated among East Asian economies.

A growing number of studies have investigated the concentration of trade within and between regions, a phenomenon described as regionalization or interregionalism. The number of regional trade agreements (RTAs), such as the single market in Europe, NAFTA, APEC, ASEAN, or Mercado Común del Sur, has been increasing steadily since the 1950s, as has the number of preferential trade agreements (PTAs) between two or more states (Mansfield & Milner, 1999).

The regionalization of trade through RTAs has been treated as both evidence for growing protectionism and increasing liberalization. The key issue is whether these agreements, which lower barriers between participants, handle trade relations with nonmembers. If they grant nonmembers equal preferential treatment, they might positively affect trade liberalization; if not, they could potentially undermine multilateral trade negotiations and global trade relations by creating exclusive trading blocs.

The Impact of Multilateral Institutions: From GATT to WTO

The dramatic growth of postwar global trade can partly be explained by looking at the multilateral institutional and regulatory framework, which has been governing global trade relations. Under the leadership of the United States in its role of a liberal hegemon, the General Agreement on Tariff and Trade (GATT) was established in 1948 to ensure that the devastating effects of protectionism during the interwar period were not repeated. GATT, which originally was to be folded into a planned International Trade Organization that was successfully negotiated but failed U.S. Senate ratification, focused primarily on the reduction of tariffs. However, its agenda later was expanded to include nontariff barriers such as import quotas, export subsidies, voluntary export restraints, and antidumping duties. After the Uruguay Round, completed in 1994, the average tariff for DCs was reduced from 6.3% to 3.8% (WTO, 1996).

The Uruguay Round also resulted in the establishment of the WTO that included the GATT, broadened the agenda of international trade negotiations by incorporating new agreements such as the General Agreement on Trade in Services, the Agreement on Trade Related Intellectual Property Rights, and the Agreement on Trade-Related Investment Measures and greatly improved the dispute settlement procedures, which had been rather weak under the GATT framework. The WTO thereby became the main international institution to address trade-related policy issues and promote free trade. As a result of all these changes, international trade has experienced far-reaching liberalization of trade barriers across the globe. IPE scholars have asked why trade liberalization endured despite the decline in U.S. trade hegemony from almost 30% of manufactured exports in 1953 to about 13% by the late 1970s. Liberal institutionalists (Keohane, 1984) and more recently constructivists (Rittberger, 1993) explain this stability with the role of regimes in upholding principles and norms that states establish and eventually internalize. Other scholars have investigated the influence of domestic factors, such as the delegation of negotiating power from legislative to executive branches of government as well as pressure from export- or trade-oriented interest groups, which lobby their respective governments for further trade liberalization (Sell, 1999). Finally, scholars have identified globalizing forces, defined as growing economic interdependence between national economies, which leave little room for politicoeconomic alternatives to an open world economy (Strange, 1994).

Future Directions

Considering the plurality of theoretical and methodological approaches as well as issue areas, any prediction on future research in the IPE of international trade must remain selective.

One strand of research focuses on the growing power position of LDCs in the multilateral trade negotiations of the WTO. Disillusioned with the unsuccessful attempts to change the structures of global trade since the 1970s through a New International Economic Order, LDCs also faced mixed consequences from the Uruguay Round, renewing discussions between liberals and historical structuralists about the effects of trade liberalization (Hoekman, 1997). Therefore, the current Doha Round has been marked by more pronounced and better organized trade diplomacy by LDCs. This often takes place in the form of country groups or blocs such as the G-20 or G-90, which are led by new trade powers like Brazil, India, or China, and challenges the traditional power relations between developed and developing countries.

A related topic has become the closer empirical investigation of the relationship among liberal markets, trade liberalization, and development, especially since the Washington Consensus, which promotes the positive effects of rapid liberalization, has increasingly been shaping

the development policies of many donor countries and UN agencies (Stiglitz, 2006).

MNCs and their impact on the structural transformation of trade have become another topic of scholarly interest. Inter- and intrafirm trade and related phenomena such as transfer prices have qualitatively changed trade, since transnational production and distribution networks represent extremely complicated structures and regulatory challenges for states and international organization, lending MNCs unprecedented structural power (Vernon, 1998).

The relationship between multilateral trade negotiations and the growing number of RTAs has also become an important research topic, as has the role of institutions for the governance of trade on the regional as well as the global level. The questions here are whether RTAs are the second-best way toward liberal trade relations or whether they generate new rifts between participants and states remaining outside these emerging trading blocs, and how institutions shape and stabilize states' expectations and help managing the IPE of trade (Cohn, 2002).

Finally, new phenomena like fair trade and the growing role of civil society actors in international trade diplomacy have come to the attention of scholarly interest. The goal of the fair trade movement is to empower producers in developing countries and conduct trade in a less exploitative and more socially responsible manner by reconnecting consumer and producer in a qualitatively new relationship on a global scale. Since the fair trade phenomenon represents alternative socioeconomic behavior that conflicts with concepts of rational action and most efficient resource allocation, mainstream rationalist IPE theory has not been able to adequately explain norm-based socioeconomic processes like fair trade (Archer & Fritsch, 2010). Fair trade is just one example for how civil society groups and organizations are increasingly impacting global governance structures and processes. Further research needs to investigate whether this participation serves the goal of improving democratic legitimacy of international institutions without negatively affecting the efficiency of international trade diplomacy.

References and Further Readings

- Abu Lughod, J. L. (1989). *Before European hegemony: The world system AD 1250–1350*. New York: Oxford University Press.
- Adler, E. (1986). Ideological guerrillas and the quest for technological autonomy: Brazil's domestic computer industry. *International Organization*, 40, 673–705.
- Alt, J., Frieden, J., Gilligan, M., Rodrik, D., & Rogowski, R. (1996). The political economy of international trade. *Comparative Political Studies*, 29, 689–717.
- Archer, C., & Fritsch, S. (2010). Global fair trade: Humanizing globalization and reintroducing the normative to international political economy. *Review of International Political Economy*, 13, 103–128.
- Ashley, R. K. (1983). Three modes of economism. *International Studies Quarterly*, 27, 463–496.

- Biersteker, T. (1993). Evolving perspectives on international political economy: Twentieth century discontinuities. *International Political Science Review*, 14, 7–33.
- Clark, C., & Bahry, D. (1983). Dependent development: A socialist variant. *International Studies Quarterly*, 27, 276–293.
- Cohn, T. C. (2002). *Governing global trade: International institutions in conflict and convergence*. Burlington, VT: Ashgate.
- Cox, R. (1983). Gramsci, hegemony and international relations: An essay in method. *Millennium: Journal of International Studies*, 12, 162–175.
- Frank, A. G. (1967). *Capitalism and underdevelopment in Latin America: Historical studies of Chile and Brazil*. New York: Monthly Review Press.
- Gereffi, G. (1983). *The pharmaceutical industry and dependency in the third world*. Princeton, NJ: Princeton University Press.
- Gerschenkron, A. (1962). *Economic backwardness in historical perspective: A book of essays*. Cambridge, MA: Harvard University Press.
- Gilpin, R. (1987). *The political economy of international relations*. Princeton, NJ: Princeton University Press.
- Grieco, J. M. (1988). Anarchy and the limits of cooperation: A realist critique of the newest liberal institutionalism. *International Organization*, 42, 485–507.
- Hamilton, A. (1966). The report on the subject of manufactures. In H. C. Syrett (Ed.), *The papers of Alexander Hamilton* (Vol. 1, pp. 230–340). New York: Columbia University Press. (Original work published 1791)
- Heckscher, E. F. (1934). *Mercantilism, vol. II*. London: Allen & Unwin.
- Held, D., & McGrew, A. (2007). *Globalization/anti globalization*. Cambridge, UK: Polity Press.
- Hilferding, R. (1981). *Finance capital: A study of the latest phase of capitalist development*. London: Routledge. (Original work published 1910)
- Hirst, P., & Thompson, G. (1999). *Globalization in question*. Cambridge, UK: Polity Press.
- Hoekman, B. R. (1997). Developing countries and the multilateral trading system after the Uruguay Round. In R. Culpeper, A. Berry, & F. Stewart (Eds.), *Global development fifty years after Bretton Woods: Essays in honour of Gerald K. Helleiner* (pp. 252–279). New York: St. Martin's Press.
- Irwin, D. (1996). Industry or class cleavages over trade policy? In R. Feenstra, G. Grossman, & D. Irwin (Eds.), *The political economy of trade policy* (pp. 53–76). Cambridge: MIT Press.
- Kaplinsky, R. (2006). *Between a rock and a hard place: Globalization, poverty, and inequality*. Cambridge, UK: Polity Press.
- Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton, NJ: Princeton University Press.
- Krugman, P. (Ed.). (1986). *Strategic trade policy and the new international economics*. Cambridge: MIT Press.
- Lenin, V. I. (1939). *Imperialism: The highest stage of capitalism*. New York: International Publishers.
- Leontieff, W. (1953). Domestic production and foreign trade: The American capital position re examined. *Proceedings of the American Philosophical Society*, 97, 332–349.
- List, F. (1916). *The national system of political economy*. London: Longmans, Green.
- Mansfield, E., & Milner, H. V. (1999). The new wave of regionalism. *International Organization*, 53, 589–627.
- Marx, K. (1990). *Capital*. London: Penguin. (Original work published 1867)
- Marx, K., & Engels, F. (1948). *The communist manifesto*. New York: International Publishers.
- Mill, J. S. (2004). *Principles of political economy*. Amherst, NY: Prometheus. (Original work published 1848)
- Milner, H. V. (1988). *Resisting protectionism*. Princeton, NJ: Princeton University Press.
- Ohlin, B. (1967). *Interregional and international trade*. Cambridge, MA: Harvard University Press. (Original work published 1933)
- Prebisch, R. (1962). The economic development of Latin America and its principal problems. *Economic Bulletin for Latin America*, 7, 1–22.
- Ricardo, D. (2006). *Principles of political economy and taxation*. New York: Cosimo. (Original work published 1817)
- Rittberger, V. (Ed.). (1993). *Regime theory and international relations*. Oxford, UK: Oxford University Press.
- Rogowski, R. (1989). *Commerce and coalitions: How trade affects domestic political alignments*. Princeton, NJ: Princeton University Press.
- Rosenau, J. N. (2007). *People count! Individuals in global politics*. Boulder, CO: Paradigm.
- Scheve, K., & Slaughter, M. (1998). *What determines individual trade policy preferences?* (NBER Working Paper No. 6531). Cambridge, MA: National Bureau of Economic Research.
- Sell, S. K. (1999). Multinational corporations as agents of change: The globalization of intellectual property rights. In A. Cutler, V. Haufler, & T. Porter (Eds.), *Private authority and international affairs* (169–198). Albany: SUNY Press.
- Smith, A. (1993). *An inquiry into the nature and causes of the wealth of nations*. Oxford, UK: Oxford University Press. (Original work published 1776)
- Stiglitz, J. E. (2006). *Making globalization work*. New York: W. W. Norton.
- Stolper, W. F., & Samuelson, P. A. (1994). Protection and real wages. *Review of Economic Studies*, 9, 58–73.
- Strange, S. (1994). *States and markets*. London: Pinter.
- United Nations Conference on Trade and Development. (1996). *World investment report 1996: Investment, trade and international policy arrangements*. New York: Author.
- United Nations Conference on Trade and Development. (2005). *Trade and development report 2005*. Geneva, Switzerland: Author.
- Vernon, R. (1998). *In the hurricane's eye: The troubled prospect of multinational enterprises*. Cambridge, MA: Harvard University Press.
- Wallerstein, I. (1984). *The politics of the world economy: The states, the movements and the civilizations*. Cambridge, UK: Cambridge University Press.
- World Trade Organization. (1996). *Annual report 1996: Trade and foreign direct investment, vol. I*. Geneva: Author.
- World Trade Organization. (2006). *World trade report 2006*. Geneva: Author.

50

NONSTATE ACTORS IN INTERNATIONAL RELATIONS

MICHAEL KUCHINSKY

Gardner-Webb University

Few could doubt that the number of organizations active in international affairs has grown sharply and that today's international relations have become more complex. These organizations go by many names—nonstate actors (NSAs), nongovernmental organizations (NGOs), private voluntary organizations (PVOs), grassroots organizations (GSOs), civil society organizations (CSOs)—sometimes used interchangeably or meant by others to signal differences of activity, structure, or purpose. Today's increasingly networked set of associations and institutions produce profound changes in the global political and economic orders.

This chapter reviews the importance of NSAs in today's international relations. Because NGOs, as an NSA, capture significant political and academic attention, they will be highlighted and discussed separately. Following an introduction that raises historical issues of NSAs and NGOs, the chapter moves through a discussion of theories for the study of NSAs, an overview of the work of some NSAs and NGOs in today's world, common critiques of the NGO sector in society, and closes with thoughts on future directions for study.

Introduction: Evolving History and Nomenclature

The decade from which we have just emerged, the 1980's, was a time of growing recognition that we live in a world in profound crisis—a world of dehumanizing poverty, collapsing

ecological systems, and deeply stressed social structures. An awareness is dawning that these are not isolated problems. Many non governmental organizations (NGOs) became increasingly aware during the 1980's that the leadership needed to deal with the underlying causes of the human tragedy was not being provided by governments. Governments and international agencies themselves came to acknowledge the ability of NGOs to do what governments have proven unable to do, i.e., to get a range of essential goods and services to the poor. Yet growing numbers of NGOs also recognized that their own efforts were too meager, and too often focused on the consequences of system failure rather than the underlying causes of this failure . . . In their attempts to deal with the reality some NGOs have sought increases in government funding to expand their service delivery capabilities. Others have questioned the nature of their more conventional roles and asked whether they may need to rethink their approaches to development actions to get at the real causes of the human suffering that motivates their action. (Korten, 1990, p. 6)

This statement, written in anticipation of the then new millennium by a scholar of state and civil society interaction, laid out many strands of conventional wisdom regarding NSAs in international relations. The argument runs as follows. A structural change occurred in global politics (the end of the cold war) that opened the door for increased NSA activity, NGO activity, or both. This opportunity widened further when past policies had not removed the specter of human tragedy and want in a world of increasingly

interconnected problems leading to a sense of impending urgency and crisis. States and international governmental organizations (IGOs) recognized their limitations in this changing world as well as the capacities of NSAs and NGOs to deliver services, giving them added legitimacy, respect, and prestige. These same organizations experienced not only the systemic transitions that the increased service possibilities provided by states and IGOs afforded them but also the equally challenging internal changes to their organizational missions and networks when confronted with global challenges and a new value-added identity. This internal set of questions challenged core assumptions of NGOs: Whose are they or whom do they represent, and what is the best means for solving global problems for the poorest people?

Nearly 20 years have elapsed since the statement was written. In between, the Soviet Union collapsed; the European Union expanded; the war on terrorism and the attacks of September 11, 2001, are nearly a decade old; climate change and globalization have become almost household words; information technology encircles the planet; and China and India have become economic giants. Even with all this, many of the aforementioned assumptions and thoughts are still in place and readily repeated by scholars and practitioners alike.

Defining which are NSAs and agreeing on terms has been difficult for international relations. One broadly inclusive text on the work of NSAs in international politics describes them as the organizations that come from civil society, the market economy, or from a political impulse, largely independent from government for funding and control, who operate across the boundaries of two or more states, and who act to affect political outcomes (Josselin & Wallace, 2001). In this way, the Roman Catholic Church, an organized crime syndicate involved in narco-terrorism, a public-policy think tank, Ford Motor Company, CARE, and even political parties that engage democratization training all fall under the umbrella of being NSAs engaged in international relations. Al Qaeda also finds a home in this definition. The breadth of this definition makes generalization difficult.

Two interrelated yet different concepts within this discussion are civil society or the civil society organization, and NGO, which is often interchanged with PVO. This is the case, for instance, for registry with the United States Agency for International Development (USAID). Civil society is the older of the terms, bringing such historical antecedents as the works of Alexis de Tocqueville and earlier voices of the Enlightenment. More current reflections of civil society embrace much of the category of the NSA while separating out market, governmental, and family institutions and organizations. Civil society generally involves some form of collective action anchored by common or shared values and interests and may be observed through formalized institutions as well as less formal and transient social movements. In this case, the civil society universe is shared by all of the street protestors and antiglobalization movements at the 1999 World Trade Organization meetings

in Seattle, as well as a labor union, a nonprofit organization, a foundation, a cooperative, or again the NGO.

Nongovernmental organization is a more self-limiting term and is often the first agent thought of as an NSA. NGOs are important since much of the research available on NSAs in international relations does not track narcotics organizations, think tanks, or local agricultural cooperatives. A vast amount of literature follows the political work and relevance of NGOs. But even here there are difficulties on account of the ubiquitous success of the NGO. The growing lexicon includes large international NGOs (INGOs), government sponsored or organized NGOs (GONGOs), and local and indigenous grassroots support organizations (GROs and GSOs). CSO is sometimes a synonym with GSO, while other NGOs may be technical aid agencies (TANGOs). Some NGOs are international, private sector, for-profit programs and consulting corporations (ICCs). The ICC is an anomaly in the discussion of NGOs but operates in the same global milieu and competes for funding and programs. In addition to institutional size and the primary level of agency (international, national, or local) in which the NGO participates, an NGO may be an operational service provider, an advocacy NGO that seeks to change public policies or educate constituencies, or a donor of resources to more local NGOs and networks.

The history of NSA involvement in international affairs is almost as complex as the identity of the actors. By tracing the registered start-ups of nonstate and nonprofit organizations, Boli and Thomas (1999) identify three global growth spurts for NSA and NGO development. These included the latter quarter of the 19th century through 1915, the years between the wars, and a steady increase in the number of agencies after 1945. Missionary societies, humanitarian assistance agencies (such as the International Red Cross), and professional and labor-related organizations dominated the first period of growth, while a number of philanthropic organizations (such as the Ford Foundation) and emergency relief organizations (such as Save the Children) began between the war years. Many of today's INGOs (Catholic Relief Services, Lutheran World Relief, Church World Service, and CARE) started at the end of World War II, leading to an explosion of NGO growth from the late 1960s onward. Though the earliest NGOs tended to be global in scope, the greatest recent growth has been in regional and grassroots NGOs that are linked by transnational networks. The growth of INGOs (or NGOs and any other NSA) is born out by the numbers—200 organizations in 1900 and nearly 6,000 by the end of the century, most which focus on economic, technical, and scientific programming or regulation.

One conclusion of this chapter is that the evolution of roles and activity for NSAs in international relations continues. These activities have intensified, diversified, and expanded even more than scholars at the end of the cold war imagined. At the same time, there has been continuity of service (humanitarian relief services or development programming), too, in keeping with the optimistic assessment of

NGOs when the post–cold war world was young. The next section grapples with the history and emerging identities of NSAs and reviews some of the discussions of these agents in international relations theory. Because of their significance, NGOs will be separated alongside their wide range of capacities and levels of activity in the global system.

Theories That Bind the NSA and NGO Community

Interest in the work of NSAs has expanded along with their numbers. Because NSAs do so many things in different spheres of activity, placing them within one theory of international relations also becomes more difficult. NSAs and NGOs complicate the theoretical environment.

Realist theories have dominated international relations, foreign policy, and security studies since the middle of the 20th century. It would be difficult to begin examining the importance of NSAs and NGOs in international relations by relying only on realist theories. The realist presupposition for systemic anarchy where states must act, often violently, to maintain their security and national interests leaves little room for NSA impact in the global order. Neorealists take more account of NSAs and NGOs in the processes of international relations. However, their presence provides new places for the state to regulate, govern, and intervene. This may be especially important since there has been uneven growth in the presence and capacities of NSAs and NGOs in today's global order (more in the global North than in the global South) that help maintain a traditional hierarchy of states and IGOs (Smith & Wiest, 2005).

International relations liberalism has traditionally described a world order that is less dependent on states to monopolize power. Liberalism makes space for transnational interaction between states and IGOs as well as major economic institutions such as MNCs, thereby redistributing power across the global system. Liberalism highlighted the importance of global norms and human rights in international relations. The theory of complex global interdependence, as discussed in another chapter, introduced a flattened hierarchy of issues beyond state security; posited a lessening of the importance of militaries for today's international relations; and developed an argument for multiple channels of political engagement to connect important actors, issues, and global agendas (Keohane & Nye, 1989). Turbulence theories expanded on the reality of a multicentric world (as opposed to a state-centered one) by highlighting multiple centers of authority, how local and subnational agents produce order and disorder, and the importance of skills and capacity development that allow more organizations and individuals to be stakeholders in international affairs. The turbulent world enmeshes macro- and microagents in a cascading process of change where states rely more heavily on private and nongovernmental resources (Rosenau, 1990).

Several theories use the regulative and dialogical relationships between states and nonstates to explain enhanced global cooperation. Regime theories reviewed how the relationships between states and international regulations produced global order and ways that nonstates helped in sustaining these international regimes. Here NSAs, and in particular the increasing presence of NGOs, play important roles in sustaining international issue-specific regimes. Theorists of epistemic communities—global communities of knowledge experts sharing insights and normative commitments from their field of expertise—focus on ways such communities influence states and public policies, especially when states must rely on these communities to better understand complex interlinkages among issues and causes (Haas, 1992).

Another set of theories that investigates the place, role, and expansion of NSAs are the works of authors using world polity, world culture, and transnational theories. Indebted to an international society and rooted in globalization studies, world polity ascribes a long-term trend of an increasing commonality of values that had its origins in European states and European global dominance from the late 18th century forward and intensified in the second half of the 20th century. World polity thinkers identify NSAs, NGOs, and the wider scientific community (epistemic communities, think tanks, and academia) as important conduits of a common world culture that relies on scientific and technologic aptitudes, increased professional organization of activity, rationalism, and individualism for its core values. NGOs become transmitters of this world polity and culture in a system headed by states and IGOs (Boli, 2005). Any evidence of an expanding world culture and an increasing number of NSAs does not remove the presence of competition and conflict with surges toward cooperation and harmony in the global order. Transnational theory builds on an international society model by incorporating NSAs into a more elaborate process for making a less hierarchical and transnational society where NSA and NGO influence augments the world of states and IGOs (Naidoo, 2006).

Constructivism contributes to the understanding of NSAs and NGOs, especially to the importance constructivism places on identities, values, and differences of meaning in the international system. Each international actor's roles are socially constructed or defined and evolve within the system of global realities and relationships. Constructivist theory accommodates the work of NSAs and NGOs because the international system is not a fixed and unchanging place. Constructivist theory is used to contextualize a wide range of activities and issues for NGOs, identify their places in the global order, and better understand their powers of persuasion, moral authority, and communicative capacity (Ahmed & Potter, 2006).

The idea that NSAs and NGOs operate within an evolving agenda of competition, cooperation, and at times conflict with the work of states leads to theories where the institutions of the state intersect with those of civil society.

Complicated relationships are created as states and their societies share characteristics of weakness and strength resulting in both nonviolent and violent change (Migdal, 1988; Ndegwa, 1996).

Discursive theories, building on the communication capacities in the nonstate sector, focus on changes in policy brought by the presence of nonstate agents in public policy debates. Using concepts such as *legacy discourses* versus *new discourses* to identify competitive meanings brought by the performance of new actors such as NGOs, discursive theory investigates how existent overlapping discourses alter the way public space is perceived and the resulting new institutions that may result. Much of the work in this area is developing and uses case study analysis to describe and assess results (Maguire & Hardy, 2006).

A useful metaphor for how NGOs work globally might be the bridge or the act of bridging. The NGO acts as a middle agent with several capacities and is therefore able to bridge across levels of society, among population groups, and between issues in order to facilitate transactions, increase networking between global and local actors, or solve problems. Thus, the NGO operates as a facilitator, negotiator, advocate, and sometimes a catalyst for program development and social change (Sikkink, 1993).

Some theories of the international political economy help to understand the operations of NSAs and NGOs. Organizations including MNCs, development NGOs, global crime syndicates, and think tanks such as the International Food Policy Research Institute all contribute to and sustain the global political economy. Susan Strange (1994) provided a complex vision for the political economy that included four interdependent structures that produce competing pressures and influences to sustain global power relationships. For Strange, the forces that structure technology, knowledge, and the world's productive capacities are highly relevant to a discussion of NSAs, especially the work of NGOs and MNCs.

And finally, the dialogue within globalization as the world's political economy provides many scholarly and popular theoretical references relevant to the work of NSAs and NGOs. The works of Thomas Friedman and Walden Bello provide two very different vantage points on globalization and the importance of NSAs in the world economy. Friedman's (2005) flat world is another form of world culture arising in our time. There could be a harmony of interests as globalization progresses. Bello's (2002) NSAs would better serve all by resisting globalization and the regulations and institutions that uphold hegemonic power.

Applications and Agency of NSAs in International Relations

NSA activities in global politics are legion. NSAs are active, systemically interacting with IGOs and regional organizations; with states, operationally providing services or advocating for public policies; and in grassroots society. The term

nonstate actor, as explored earlier, describes a broad set of actors whether providing productive services or engaged in negative and violent behavior. NGOs are the largest component of the NSA community, and their work is highlighted in the next section. Few issues areas do not attract the attention of NSAs or NGOs. States and IGOs have increased their use of NSAs, and texts on globalization (Josselin & Wallace, 2001) speak to their increased importance for the global economy for building capacities and skills and sharing information and technology to the world's peoples. This section begins by addressing selective activities of lesser-known NSAs, in particular the roles of think tanks, religions, and transnational criminal actors before addressing the work of NGOs. The work of multinational corporations and international terrorist groups also fit the NSA categories, but these are covered more extensively in other chapters.

Although older think tanks such as the Brookings Institution in the United States began before World War II, and other larger think tanks started after the war with the establishment of the UN system and the cold war's bipolar structure, the largest growth in think tanks emerged after 1970 because of an emerging crisis or in support of an increasingly dominant and global neoliberal economic model. Think tanks provide research, program assessments, expertise, and consultative services on social policy issues, generally to IGOs and governments, but are also sought out by media outlets and other NGOs. Their research is intended for the development of public policy while interacting globally with other think tanks and NSAs. Think tanks may initially have started as centers for general study, innovation, and objective knowledge, but today's think tanks are more diverse. Think tanks may be highly specialized in their focus (such as the NGO-embedded Bread for the World Institute), ideologically conservative (the CATO Institute) or left of center (Institute for Agriculture, Trade, and Development), located in the developing world (Focus on the Global South) for the purposes of challenging Western knowledge and the policy prescriptions of IGOs, linked to a national government (United States Institute for Peace), contribute expertise to the public policy concerns of a region (the Southern African Research and Documentation Centre), or to the global needs of the United Nations (Consultative Group on International Agricultural Research).

In addition to creating knowledge, an increasingly important component of think tanks in an interdependent world is their ability to transfer ideas and schemes among recipients, thereby increasing their global outreach. Formal and informal policy networks, epistemic communities, and international conferences aid in this activity.

Think tanks transfer ideas and ideologies, policy proposals and justifications, personnel and expertise, as well as documents for policy discussion and exchange. They are a pool of knowledge, resources, and expertise concerning the policies of other countries, localities, or regions. They help transfer the intellectual matter that underpins policies. Institutes can provide the rhetoric, the language and the scholarly discourse to

give substance and legitimacy to preferred policy options. Such arguments can then be used by other actors in governmental or party political debates involving policy transfer. (Stone, 2001, p. 353)

Think tanks accentuate influence by positioning themselves in relation to significant global actors, providing the language and examples that become established assumptions for solving global problems, and through their evaluation of policy prescriptions and programs.

Religious organizations can also be NSAs. Sometimes identified as “moral-entrepreneurs” (Colonos, 2001) who are able to mobilize people and organizations for ethical positions of support or protest or viewed through the actions taken by religious elites, religions have been long-time participants in international relations. With their leaders, constituencies, principles and ideological positions, and physical and financial resources, religions may operate like any organized political interest group. With increasing frequency, research devoted to religion(s) in international relations, the political economy, or foreign policy (Cheng & Brown, 2006; Hanson, 2006) recognize the importance of religious communities and leaders in international affairs. Individual case studies of religious activism on issues of conflict, social justice, the economy, women and minorities, or human rights are still more common. Many religions maintain offices of governmental affairs in national capitals or in IGO headquarter cities (New York or Geneva) that provide information to governments while simultaneously advocating their policy positions.

The increasing importance of NSAs in a globalizing world has not always been progressive. The same opportunities that promote NGOs, think tanks, or MNCs to operate—more porous boundaries, increased socioeconomic transactions, and the importance of information technology—assist organizations that may be considered less desirable in today’s world. Much has been written about terrorist groups and violent religious fundamentalist organizations since the tragic events of September 11, 2001, and research is filling gaps on current terrorist analysis and the activities of organized religious fundamentalists. Less well covered is transnational organized crime.

An exception focuses on the structure of narcotics trafficking, narcotics cartels, the global regulations that affect them, and the international system of agencies required for the financing and distribution of narcotics products. In this case, the illegal narcotics industry was analyzed using an international commodity model not unlike those monitoring any MNC product, with one major difference—financial lack of transparency and money laundering (Mares, 2006). Similar research is now becoming available on aspects of global sex trafficking.

Nongovernmental Organization Actions

By sheer numbers, NGOs dominate the world of NSAs. Their history is marked by expansion—expanding numbers,

responsibilities, issues, and evolving practices due to their successful or failed practices. NGOs operate with the UN, among state governments, in grassroots communities, and transnationally. They include large INGOs whose programs are in dozens of countries, employing thousands of workers, funded by numerous (states, foundations, or IGOs) donors, and whose budgets run into the millions of dollars. NGOs can be as small as a single project in a village staffed by few employees and supported by one donor. NGOs divide functionally among agencies that provide humanitarian emergency services, those that are operational in program delivery for economic development, and those that focus on influencing public policy or broadening human rights applications through advocacy and education. Many (such as Church World Service) provide all three services—humanitarian relief, operational and technical services, and public policy work. When viewed holistically, NGO presence and program application in international affairs are as impressive as they are confounding for easy analysis.

Some of the earliest NGOs were extensions of religious communities (such as the Mennonite Central Committee and the Mennonite church, Catholic Relief Services, or Lutheran World Relief) offering voluntary humanitarian assistance to war or famine victims, refugees, or displaced persons and originating around a major war of the 20th century (Lutz, 1994). They never relinquished these services even though their programs would expand across time. These same early NGOs moved their attention away from postwar Europe to the developing and newly independent states of Asia and Africa as colonialism ended. Religious agency still commands a significant part in NGO activities, but these, too, have diversified beyond any European, North American, or Judeo-Christian roots (Bond, 2004; Sullivan, 1994).

The 1990s onward have been decades of growth for NGOs and GROs. Several events and trends converged to catapult NGOs into greater prominence. These included the 1992 UN Conference on Environment and Development in Rio de Janeiro, where thousands of NGOs came and showed the strength of transnational civil society and the end of the cold war, general disillusionment with World Bank and International Monetary Fund structural adjustment programs of the 1980s, a return to poverty alleviation models, increased interest in democratization by civil society, a diminishing pool of foreign assistance, and the greater willingness among states to involve NGOs in issues where they had been less prominent.

Development NGOs turned to interventions that would increase indigenous participation in sustainable development programs by increasing the capacities of local GROs. Complaints that this had not been achieved after decades of modernization-directed development assistance were the poignant cry of poor people from around the world (Narayan, 2000).

Several directions consumed donor NGOs in their efforts to increase participation, skills, and social capacity through sustainable grassroots networks as the best means

for poverty alleviation. The first was that agriculture and the performance of smallholder farmers needed sustained technical improvements, access to credit and markets, and cooperative arrangements that could assist against the problems of access and scale in the global economy. NGOs would target women—another traditionally ignored population for state-centric industrialization projects—especially in programs of health care and literacy, agriculture, and human rights (Edwards & Gaventa, 2001; Kevane, 2004). The success stories of the Grameen Bank in Bangladesh raised interest for microfinance, credit, and enterprise programs. And last, NGO activism in environmental sustainability has influenced policy agendas; provided research; monitored treaties; and encouraged conservation, sustainable land use, and appropriate technology (Ahmed & Potter, 2006).

Global security became a new investment area for NGO activity (Mbabazi & Shaw, 2000), prompting inter-NGO dialogue on how to resolve security challenges while at the same time creating safe spaces for NGO personnel (Avant, 2007). As service providers, NGOs would remove forgotten landmines and become involved in the reintegration and psychosocial trauma realities of returning child soldiers and the village victims of civil conflicts (Kumar, 1997). NGOs led global advocacy and education during the highly successful international campaign to ban landmines while others became active to mediate violent conflict (Yarrow, 1978).

Critical Issues of NGO Performance

Since much has been written praising NGO performance and capacity (Korten, 1990), it is not surprising that they have become open for critical evaluation. The issues include accountability, representation, legitimacy, and in some cases their politicization.

Some debates have been longstanding within the NGO assessment literature. An early voice from the 1970s preceded the future growth of NGOs. The analyst used a rational choice methodology to examine how multiple influences (mission statements, boards, agency worldview, leaders, and audience) impact choices of agency behavior. The research suggested that internal concerns (such as leadership, board influence, budget, and funding) mattered more than mission beliefs and service audience. The research uncovered that NGO advocacy was often directed toward the maintenance concerns of the agency. The early critique offered a comparison on whether NGOs were more idealistic or interested in income maximization and whether they tended to be active in relief programs or liberation. Conclusions favored income maximization and the relief agenda over idealism and long-term sustainability (Lisner, 1977).

Obviously, NGO roles and information have expanded. But the early study elicited questions regarding the legitimacy of NGO claims, whether the NGO is accountable to its

mission, and whether they represented their internal constituencies and funders more or spoke for the poorest of the poor. This latter criticism over linkage between donors and the poor was significant since it reflected on the basic assumption of the NGO's place in the global system (Carroll, 1992).

A thorough examination of relationships among governments, donors, and NGOs was necessary, especially following the cold war and the expanded presence of NSAs and NGOs. One important set of cases of Western and non-Western NGOs concluded that donor influence was greater on NGOs than vice versa and that NGOs would have to return to their roots if they were to maintain their effectiveness in the fields of poverty alleviation and development. The conclusions reflected the concern that the supply side of new resources did not always match well with a less articulated demand side of needs and capacities. Though not without successes, a cautious conclusion for the northern NGO that became increasingly linked to governments and donors was the necessity to understand the risks of getting larger, the accountability issues between the NGO and grassroots CSOs, and the NGO's purpose as an intermediary (Hulme & Edwards, 1997).

Accountability and legitimacy are inseparable for NGOs. What they do, who they are, and the norms they espouse intersect with their various stakeholders. They are vertically accountable to their funders and horizontally accountable to their mission. Thus, NGOs must be measured by their performance and the meaning they give to performance criteria. Performance accountability gets related to the transparency of the agency, the cost-effectiveness of the programming, the ability to reach the world's poorest, the ability to increase participation and assist in global democratization, and the empowerment of local capacities in the developing world, among others. Discussions about accountability often center on the functional concerns of how funds are used and short-term impacts sustained by the NGO's program, while a more interpreted and strategic accountability relates to long-term objectives, any enhanced on-the-ground capacities, as well as impact on wider development-actor relationships. Functional and strategic criteria are both essential for intermediary NGOs that connect systemic agents (states, donors, and IGOs) and the stakeholders of grassroots communities through their normative claims.

NGOs have engaged in a spirited dialogue about improving accountability. Enhanced accountability might be achieved by using learning assessments, since the process of development can never be fully achieved and is unique to every set of programs and actors; the use of social audits; the formation of citizen oversight groups within programs; or regular forums for dialogue that reform and create better international institutions (Ebrahim & Weisband, 2007; Edwards & Hulme, 1996).

A last set of issues that have dogged NGOs has been questions over their neutrality and politicization. Neutrality has been essential for humanitarian service organizations

such as the International Red Cross during and after violent conflict. Clearly, an organization may express norms (such as use the UN Universal Declaration of Human Rights), maintain a privileged position of the poor in discussions about economic growth, or choose behaviors (participate in global advocacy campaigns) that may shift their work toward favored standards or constituencies (and away from a strict neutrality). NGOs that proclaim neutrality and yet are committed to specific ideological claims are criticized. One such study reviewed the programs of a respected development NGO, Oxfam UK, and provided a history of Oxfam shifting positions about sustainable development during the *Ujamaa* campaigns of then Tanzanian President Julius Nyerere. Citing policy statements, leader positions, and campaigns, the research argued that a bias toward Nyerere's views and an ongoing Tanzanian project became Oxfam's ideology for programs worldwide (Jennings, 2002). More acerbic was David Rieff's (2002) commentary arguing that NGOs lose their usefulness by abandoning neutrality and that the cause of independent humanitarianism standing outside of political processes had been irreparably harmed for the NGO community, world society, and people in need.

Policy Implications

When so many different political actors work in overlapping issues, serve multiple constituencies, and operate at every social level and everywhere, policy coordination has been and will continue to be a concern. How does one bring about institutional and programmatic efficiencies and leverage resources and agency interrelationships at the same time? What institutions need to be built or reformed to provide and monitor economic and political development, encourage trade while overseeing global financial flows, or contain conflicts enmeshed with issues of identity that spill across boundaries? Should new regional and global institutions be crafted, or should informal and decentralized approaches dominate?

The discussions on challenges to NGOs brought out the concerns of legitimacy, accountability, and the representative quality of the agencies. The fact that this assessment dialogue was taking place in house (a form of self-regulation) might be applauded, while others would challenge the objectivity of the results. Similar intra- and interagency policy questions can be raised about the performance of MNCs and think tanks. Where studies emerge about illegal NSAs, such as worldwide narcotics trafficking, current policy cases focus on issues of production or consumption and degrees of punishment or tolerance, while global dialogues about results are less prevalent.

The collaborations between different NGOs are just one of many emerging partnership models in the evolving work of NSAs. The discussions in the United States about faith-based initiatives and agencies indicated how heated and ambiguous the nature of these emerging partnerships

might be. In recent years, these collaborating partnerships have included one of the world's largest bilateral development organizations, United States Agency for International Development, with individual religious communities working overseas on local development projects. Wondering about the policies required to monitor the collaboration and assessment of thousands of small organizations engaged in the global mission of international development that do not hinder innovation, entrepreneurship, human rights, and privacy concerns stretches the imagination even more.

Future Directions

In many ways, future research on NSAs of all types, and NGOs in particular, has never been more diverse and fertile. Several questions have been raised along the way where more work should be done—on NSAs and international relations theory or the increased importance of assessment and comparative studies of NSA and NGO accountability and legitimacy after they have become a growth factor in international politics. Organizations that were overlooked in the past by international relations, such as religious organizations, other cultural identity agencies, even the place and roles of individuals with significantly demanded skills, require further research. The efficacy of NGOs in issues traditionally handled by states or IGOs—such as conflict, security, and peace building—continues to demand attention even as definitions of security and threat analysis evolve. Work on transnational relations and the global networks that tie organizations together across state and systemic boundaries to pursue common causes has begun and provides ample room for growth. As the work of NSAs evolves and diversifies, so do the partnerships that bring them together. Understanding the multiple collaborations possible between public, nonprofit, and private sector actors in the ongoing needs of a public good, such as global environmental protections and policies, is just one blended area of research that challenges available models and concepts to analyze the collaborations' effect. These thoughts suggest that understanding the work of NSAs will demand greater attention to interdisciplinary approaches and the hybridity of analysis and theoretical modeling that makes the modern world, perhaps, postmodern.

Conclusion

Though NSAs have been involved in international relations for many years, they have seldom been as involved or entrusted with the range and depth of the world's challenges as they have today. They provide resources, employment, and improve social capacities while engaged as donors, partners, operational service providers, public policy advocates, and to a lesser degree, violent, criminal

components of a global civil society. They have had remarkable resilience and an ability to evolve into new or different organizations, in pursuit of changing missions. NSAs will continue to be vital and necessary agents among global systems, in collaboration with states, among grassroots communities, and linked together as transnational networks forming triangular relationships with sovereign states and IGOs yet blending with local communities and peoples.

References and Further Readings

- Ahmed, S., & Potter, D. M. (2006). *NGOs and international politics*. Bloomfield, CT: Kumarian Press.
- Avant, D. (2007). NGOs, corporations, and security transformation in Africa. *International Relations*, 21, 142–161.
- Batliwala, S., & Brown, L. D. (Eds.). (2006). *Transnational civil society: An introduction*. Bloomfield, CT: Kumarian Press.
- Bello, W. (2002). *Deglobalization: Ideas for a new world economy*. London: Zed Books.
- Boli, J. (2005). Contemporary developments in world culture. *International Journal of Comparative Sociology*, 46, 383–404.
- Boli, J., & Thomas, G. M. (Eds.). (1999). *Constructing world culture: International non governmental organizations since 1875*. Stanford, CA: Stanford University Press.
- Bond, G. D. (2004). *Buddhism at work: Community development, social empowerment, and the Sarvodaya movement*. Bloomfield, CT: Kumarian Press.
- Carroll, T. F. (1992). *Intermediary NGOs: The supporting link in grassroots development*. West Hartford, CT: Kumarian Press.
- Cheng, T. J., & Brown, D. A. (Eds.). (2006). *Religious organizations and democratization: Case studies from contemporary Asia*. Armonk, NY: M. E. Sharpe.
- Colonos, A. (2001). Non state actors as moral entrepreneurs: A transnational perspective on ethical networks. In D. Josselin & W. Wallace (Eds.), *Non state actors in world politics* (pp. 76–89). New York: Palgrave.
- Ebrahim, A., & Weisband, E. (2007). *Global accountabilities: Participation, pluralism, and public ethics*. New York: Cambridge University Press.
- Edwards, M., & Gaventa, J. (Eds.). (2001). *Global citizen action*. Boulder, CO: Lynne Rienner.
- Edwards, M., & Hulme, D. (Eds.). (1996). *Beyond the magic bullet: NGO performance accountability in the post cold war world*. West Hartford, CT: Kumarian Press.
- Friedman, T. (2005). *The world is flat: A brief history of the 21st century*. New York: Farrar, Strauss & Giroux.
- Haas, P. (1992). Introduction: Epistemic communities and international policy coordination. *International Organization*, 46, 1–35.
- Hanson, E. (2006). *Religion and politics in the international system today*. New York: Cambridge University Press.
- Hulme, D., & Edwards, M. (Eds.). (1997). *NGO, states, and donors: Too close for comfort?* London: Macmillan.
- Jennings, M. (2002). Almost an Oxfam in itself: Oxfam, Ujamaa, and development in Tanzania. *African Affairs*, 101, 509–530.
- Josselin, D., & Wallace, W. (2001). *Non state actors in world politics*. New York: Palgrave.
- Keohane, R. O., & Nye, J. S. (1989). *Power and interdependence* (2nd ed.). New York: HarperCollins.
- Kevane, M. (2004). *Women and development in Africa: How gender works*. Boulder, CO: Lynne Rienner.
- Korten, D. C. (1990). *Getting to the 21st century: Voluntary action and the global agenda*. West Hartford, CT: Kumarian Press.
- Kumar, K. (Ed.). (1997). *Rebuilding societies after civil wars: Critical roles for international assistance*. Boulder, CO: Lynne Rienner.
- Lisner, J. (1977). *The politics of altruism: Study of the political behavior of voluntary development agencies*. Geneva, Switzerland: Lutheran World Federation.
- Lutz, C. P. (1994). *Loving neighbors far and near: U.S. Lutherans respond to a hungry world*. Minneapolis, MN: Augsburg.
- Maguire, S., & Hardy, C. (2006). The emergence of new institutions: A discursive perspective. *Organization Studies*, 27, 7–29.
- Mares, D. (2006). *Drug wars and coffee houses: The political economy of the international drug trade*. Washington, DC: Congressional Quarterly Press.
- Mbabazi, P., & Shaw, T. (2000). NGOs and peacebuilding in the great lakes region of Africa: States, civil societies, and companies in the new millennium. In D. Lewis & T. Wallace (Eds.), *New roles and relevance: Development NGOs and the challenge of change* (pp. 189–198). Bloomfield, CT: Kumarian Press.
- Migdal, J. S. (1988). *Strong societies and weak states: State society relations and state capabilities in the third world*. Princeton, NJ: Princeton University Press.
- Naidoo, K. (2006). Claiming global power: Transnational civil society and global governance. In S. Batliwala & L. D. Brown (Eds.), *Transnational civil society: An introduction* (pp. 51–64). Bloomfield, CT: Kumarian Press.
- Narayan, D. (2000). *Can anyone hear us? Voices of the poor*. New York: Oxford University Press.
- Ndegwa, S. (1996). *The two faces of civil society*. West Hartford, CT: Kumarian Press.
- Rieff, D. (2002). *A bed for the night: Humanitarianism in crisis*. New York: Simon & Shuster.
- Rosenau, J. N. (1990). *Turbulence in world politics: A theory of continuity and change*. Princeton, NJ: Princeton University Press.
- Sikkink, K. (1993). Human rights, principled issue networks, and sovereignty in Latin America. *International Organization*, 47, 411–431.
- Smith, J., & Wiest, D. (2005). The uneven geography of global civil society: National and global influences on transnational association. *Social Forces*, 84, 621–652.
- Stone, D. (2001). Think tanks, global lesson drawing, and networking social policy ideas. *Global Social Policy*, 1, 338–360.
- Strange, S. (1994). *States and markets* (2nd ed.). London: Pinter.
- Sullivan, D. J. (1994). *Private voluntary organizations in Egypt: Islamic development, private initiative, and state control*. Gainesville: University Press of Florida.
- Yarrow, C. H. M. (1978). *Quaker experiences in international conciliation*. New Haven, CT: Yale University Press.

51

INTERNATIONAL ORGANIZATIONS AND REGIMES

BYUNGWON WOO

The Ohio State University

The study of international organizations and regimes has become increasingly popular over the past three decades. This recent resurgence of the study of international organizations is very distinct from earlier studies of international organizations in several regards. First, departing from the study of legal principles and formal organizational structures of international organizations in the earlier era, the study of international organizations has become more social scientific, with strong theoretical developments and accompanying empirical examinations of the theoretical advances. Many recent studies attempt to provide general explanations for creations, roles, effects, effectiveness, and other institutionalized features of international organizations and try to demonstrate how these general explanations hold through rigorous empirical testing. Second, along with continuing the earlier practice of investigating a single international organization in a study, recent studies start to tackle universal issues and ask questions about features present in and applicable to a group of international organizations. Examples are numerous, including voting rules, membership size, and degree of independence and legalization of international organizations. Third, the scope of the study of international organizations has been greatly broadened. In its earlier years, the study of international organizations was limited to a few prominent formal international organizations such as the United Nations and the European Union (or its earlier version, the European Community).

Yet there are numerous other formal and informal international arrangements that guide states' behaviors, ranging from formal international arrangements to informal yet widely accepted international norms. Thus, while formal international organizations still remain important subjects and the studies investigating these international organizations ever increase as a number of new international organizations have been created in the past few decades, other international arrangements, referred to as either international regimes or international institutions, have also been brought into the field of international organizations and become subjects of inquiry in the study of international organizations.

This chapter reviews this recent development of the study of international organizations over the past three decades. The first section starts with the formal definitions of and distinctions among international organizations, international institutions, and international regimes. The next two sections focus on theoretical debates among the three major theoretical orientations in mainstream international relations. The second section introduces theories of international regimes and discusses important debates regarding the creation and functions of international organizations. It highlights the important agreements and disagreements among the three main theoretical paradigms in the discipline of international relations—realism, neoliberal institutionalism, and constructivism—and ends with a discussion about the ways in which one might bring the competing theories together for a more comprehensive

understanding of the creation and functions of international organizations. The third section visits more recent scholarship, including rational institutional design and delegation literatures. The concluding section recaps the chapter and discusses a few potential future research agendas in the study of international relations.

Definitions

Although the terms *international organizations*, *international institutions*, and *international regimes* are often used interchangeably, their precise definitions are slightly different from one another. Thus, before delving into the discussion of reviewing the extant studies of different aspects of international organizations and regimes, the formal definitions for each term are introduced, and differences among the three terms are briefly highlighted in this section.

International Institutions

Although scholars adopt slightly different definitions for international institutions, Keohane (1988) defines international institutions as persistent and connected sets of rules that prescribe behavioral roles, constrain activity, and shape expectations. These rules may be formal and explicit or informal and implicit.

Treaties between more than one sovereign state, including international agreements, covenants, conventions, and protocols, are all good examples of formal international institutions. They are signed and ratified by more than one state and generally guide and constrain participating states' behaviors and shape their expectations about future behaviors of each other. For instance, the Universal Declaration of Human Rights prescribes what to and what not to do in respecting citizens' basic rights within states' territories. Similarly, bilateral or multilateral defense treaties between allies prescribe obligatory actions in conflict scenarios and thus shape expectations of alliance behaviors in case of potential military conflicts. With growing international economic transactions, a growing number of multilateral and bilateral free trade agreements have been signed between states in recent decades so that states can regulate the trade practices of signed parties.

Along with formal international institutions signed and ratified by states, there are other informal international institutions. The most basic international institution is the principle of sovereignty. Sovereign states claim rights that other international entities cannot. Sovereign states have exclusive rights over their territories and people; thus, in principle, their domestic matters should not be interrupted by other states. In addition, sovereign states are treated equally, at least in a legal sense, regardless of their economic wealth, size of population, or military might. The principle of sovereignty also serves as the basis of other

formal and informal institutions, and the evolution of the principle of sovereignty often leads to changes in other international institutions.

International Regimes

International regimes are defined as principles, norms, rules, and decision-making procedures around which actor expectations converge in a given issue area. Principles and norms provide the basic defining characteristics of a regime, and there may be many rules and decision-making procedures that are consistent with the same principles and norms (Krasner, 1983).

The term *international regime* has generally been used to refer to rules and norms within a particular issue area. For instance, an international nuclear nonproliferation regime has been formed to manage and limit both horizontal and vertical nuclear proliferation in the world. Within the nuclear nonproliferation regime, there are many agencies and international treaties that perform detailed functions and assign specific rules. Under the defining principle of nonproliferation of nuclear weapons beyond the five countries that had tested nuclear weapons by the beginning of 1960s, there are many specific rules and procedures that constitute the international nuclear nonproliferation regime, such as safeguarded nuclear facility inspection procedures by the International Atomic Energy Agency and the comprehensive nuclear-test-ban treaty.

Since international regimes are issue specific, it is important to point out that general-purpose international organizations such as the United Nations or the European Union are not considered as regimes (Martin & Simmons, 2001). Instead, organizations like the United Nations and the European Union participate in a number of issue-specific regimes. For instance, the United Nations encompasses multiple regimes, with agencies involved in one or more regimes, such as nuclear nonproliferation, peacekeeping, economic development, global health management, human rights, and environment protections.

International Organizations

Simply put, international organizations are the formal embodiment of the international institutions and regimes discussed previously (Martin & Simmons, 2001). They are housed in buildings, employ international civil servants and bureaucrats, and have nontrivial budgets for their operations. They are usually created by the international treaties that serve as the basis for their continuing operations.

According to the *Yearbook of International Organizations*, published by the Union of International Associations (2007), there were 242 international organizations as of 2007. The number of international organizations peaked in late 1980s, when the Soviet bloc and the Western bloc maintained their own international organizations. The number has decreased and stabilized since the collapse of the Soviet Union. The

international organizations listed in the yearbook include not only well-known international organizations like the United Nations, the European Union, the World Bank, and the World Trade Organization (WTO) but also lesser-known organizations such as the International Coffee Organization and the International Whaling Commission.

Only sovereign states can be members of international organizations, and this differentiates international organizations from other international entities whose members include individuals and groups of individuals from more than one country. These international entities are usually referred to as international nongovernmental organizations. These organizations have memberships and activities in more than one country, and most of them are not for profit.

There are a very large number of nongovernmental international organizations across diverse issue areas. The *Yearbook of International Organizations* also keeps track of these nongovernmental international organizations. As of 2007, there are more than 7,500 nongovernmental international organizations, and the number is still growing. A few well-known nongovernmental international organizations include Greenpeace, Oxfam, International Committee of the Red Cross, Amnesty International, and Doctors Without Borders. These organizations often play critical, often complementary, roles to the activities of formal international organizations in a number of important international issues such as humanitarian assistance, environmental protection and pollution monitoring, human rights monitoring, developmental assistance, and conflict resolution. Nongovernmental international organizations can often work more expediently when needed, because they are relatively freer from internal conflicts and disagreements among their members compared with international organizations where disagreements among their member states often hinder effective and expedite action taking. One reason for this is because states' interests often vary widely, while participants in nongovernmental international organizations share a common goal. Thus, nongovernmental international organizations have become increasingly active in international relations. However, little scholarly attention has been paid to study them thus far; hence, the following sections focus only on international organizations and international regimes established by sovereign states.

Theories of International Organizations and Regimes

For the past few decades, the debates between two prominent international relations theoretical orientations—neorealism and neoliberal institutionalism—have centered on the prospect of cooperation and the roles of international organizations in facilitating and promoting cooperation. In particular, a series of research studies in the tradition of neoliberal institutionalism have made major advances in our understanding of why states create international organizations and what roles international organizations play in facilitating international cooperation. On the other hand, realism, the

research tradition emphasizing states' power in explaining international political phenomena, provides reasons why international organizations are biased toward more powerful countries' interests and play only limited roles in facilitating cooperation among states, if at all.

Since the early 1990s, the new approach in international relations emphasizing norms, knowledge, culture, and ideas, generally referred to as constructivism, emerged and rapidly established itself as a main theoretical orientation in international relations. Constructivism encompasses a very broad spectrum of scholarship that commonly focuses on the role of ideas. The spectrum ranges from studies in the scientific positivist tradition emphasizing nonmaterial factors to studies in postmodernist tradition. Studies of international organizations in the constructivist tradition have provided powerful critiques to and pointed out limitations of studies in realist and neoliberal institutionalist traditions since then.

Neoliberal Institutionalism

Studies in the neoliberal institutionalist tradition start with the premise that states act rationally and pursue the best possible strategy given a situation. Neoliberal institutionalists also assume that the anarchic nature of international system constrains the ways in which states interact and makes it difficult for states to cooperate. But neoliberal institutionalists argue that states can often overcome the constraints of anarchy, achieve cooperation, and reach mutually beneficial outcomes by creating international organizations.

The international system lacks an authority that sets, monitors, and enforces rules. This is a very distinct characteristic of international politics as compared with its domestic counterpart where laws are established, implemented, and enforced by legislative, executive, and judicial branches of a government. This anarchic nature of the international system and the reason it is difficult to achieve cooperation under anarchy is often illustrated with the prisoner's dilemma.

A generic version of the prisoner's dilemma features two actors who need to decide, without communicating with each other, whether they want to cooperate or not cooperate. The four preferences over the possible outcomes are the same for both actors and are given in order, from most preferred to least preferred: (1) not cooperating when the other actor cooperates, (2) cooperating when the other actor cooperates, (3) not cooperating when the other actor does not cooperate, and (4) cooperating when the other actor does not cooperate.

If both actors act rationally, in this unique equilibrium, both actors decide not to cooperate, regardless of one's expectation of the other actor's behavior; thus, they end up with the third-preferred outcome: Both do not cooperate. This is because when an actor expects the other to cooperate, it is in the actor's best interest not to cooperate since the actor prefers Outcome 1 to Outcome 2. Similarly, when an actor expects the other not to cooperate, it is in the actor's

best interest not to cooperate since the actor prefers Outcome 3 to Outcome 4. Thus, if both actors play rationally, then they end up with the respective third-best alternative. This is less than what both could have achieved if they had both managed to cooperate. This dilemma of reaching the collectively suboptimal outcome when individually acting rationally poses the central puzzle in achieving international cooperation for neoliberal institutionalists.

Consider the dilemma in a provision of international public goods context. There are a few international public goods such as fish stock in a blue ocean, clean air, and the ozone layer. Suppose that states in the world need to decide whether to cooperate or not to cooperate to contribute to the supply of the international public good, say clean air, by reducing their current levels of carbon dioxide emissions without knowing if other states will reciprocate. If all states cooperate and reduce emission levels, all can enjoy the international public goods, with each bearing due costs; thus, they can reach the mutually beneficial outcome. Yet if one defects while the others cooperate, one can enjoy what the others contribute, without paying the due cost and perhaps fuel more rapid economic growth. In contrast, when one cooperates while the others defect, one pays its own cost while the others free ride or worse, the international public good is not provided. Finally, when all states defect, no public good is provided and everyone is worse off than when all states cooperate and enjoy the public good. Following the same reasoning, what is the best strategy for an individual state? When a state believes that the others would provide public goods, it is in the state's best interest to defect, since the state can free ride and enjoy the public good as compared to paying the cost and enjoy the public good. Similarly, if the state believes that the others will defect, the best strategy for the state is to defect as well, since the state would not want to be the only one to contribute to the public good provision while others do not reciprocate. Thus, regardless of one's belief regarding the others' behavior, it is in the state's best interest not to contribute to the public good provision. And if every state in the international system reasons this way, then no international public good is provided and every state is worse off and suffers consequences of global warming, as opposed to the hypothetical situation where all states contribute and enjoy clean air.

The problem of the noncooperation is not that states are malevolent. Rather, the suboptimal outcome of noncooperation stems from states acting individually rationally, and states choose not to cooperate because they fear that they may end up being cheated by the other states. Specifically, there are three reasons that states decide to not to cooperate. First, states decide not to cooperate because there is no behavioral standard in the first place. Second, even if there exist certain rules that states agree to abide by, there is no monitoring mechanism to monitor each other's compliance. Third, there is no international police or judiciary body to punish those who do not comply with the rules.

Then what is the solution? For neoliberal institutionalists, international organizations exist to solve these kinds of

problems inherent in the logic of anarchy. The early neoliberal institutionalist argument is articulated in Robert Keohane's (1984) seminal book, *After Hegemony*, and a series of scholarly articles published in the leading scholarly journal, *International Organization*, since then have elaborated on how international institutions can facilitate mutually beneficial cooperation. Neoliberal institutionalists argue that institutions can facilitate mutual cooperation by reducing uncertainty for engaged parties and reducing transaction costs. Specifically, international institutions provide the basic rules in a given area on which states can judge legitimacy of each other's behavior. Moreover, some international institutions have built-in monitoring mechanisms. Reduced uncertainty about the behavioral standards and each other's potential violations lets states be more willing to cooperate. This institution-as-uncertainty-reducer argument is further supported by the studies that suggest that even without a central, enforcing authority regulating transactions, mere provision of information can significantly enhance the prospect of cooperation (e.g., Milgrom, North, & Weingast, 1990).

Institutions also create a more amenable environment for cooperation among participating parties by allowing them continued interaction (Axelrod, 1984). International organizations let states interact more regularly and repeatedly in multiple issue areas, and the enhanced prospect for iterated interactions accommodates cooperation among countries. When states expect to frequently interact with each other in the future, states can reciprocate each other's behavior in the future interactions. This expectation of future interactions reduces the temptation to cheat in the current interaction and reduces the fear of being cheated since one can retaliate in the future. Thus, international organizations, by prolonging what is termed the *shadow of future*, also can promote cooperation (Axelrod, 1984; Oye, 1986).

Stein (1982) categorizes problems in international cooperation into those of coordination and those of collaboration. Coordination problems are the problems that do not require active monitoring and enforcement operations of international organizations because they are self-enforcing—there is no incentive for parties to defect once they make an agreement. In this case, international organizations need only to provide a focal point that each party's expectation can converge around. For instance, international aviation requires a common medium to communicate. Without the common communication medium, one cannot guarantee smooth communication between American pilots flying over China and Chinese air traffic controllers or vice versa. And if there are communication problems, then the safety of interstate aviation cannot be ensured. Neoliberal institutionalists would argue that states established the International Civil Aviation Organization and set the official language for international aviation as English, in order to solve the coordination problem of international aviation. Once the international organization sets the rules and presents the rules to the states, then the rules are self-enforcing. That is, there is little incentive for participating states to

cheat and instead speak French or Chinese for international aviation communication since the set rules are mutually desirable. Most of the functional international organizations, such as the Universal Postal Union, whose main purpose is to set a series of rules for smoother international operations, would fit this category.

For collaboration problems, institutions need to perform other roles to promote mutually beneficial cooperation in addition to provide rules. A typical collaboration situation resembles the prisoner's dilemma. In international relations, trade agreements among states and public goods regimes such as the environmental regime fit into this category. These situations are characterized by lingering incentives to cheat even after reaching mutually desired outcomes, since cheating would allow free riding and provide short-term benefits. Thus, international organizations created for solving collaboration problems need to provide additional monitoring and enforcement mechanisms. The International Atomic Energy Agency and the WTO have indeed monitoring and enforcement procedures that states can rely on. In the WTO case, when states suspect that there are certain countries that do not comply with the trade rules, states can ask to establish a dispute settlement panel and initiate a legal case against those countries.

Building on the earlier work of Keohane (1984) and responding to the realists' claim that institutions exert little independent influence in world politics, Keohane and Martin (2003) further elaborate on the neoliberal institutionalist position. In doing so, they provide three additional reasons that why international institutions do matter. First, when there are many possible agreeable cooperation points, existing institutions often play as a focal point by providing guidance around which states' behaviors can converge. Second, Keohane and Martin argue that institutions, once created, often have their own lives and become "sticky." The International Monetary Fund (IMF) is a good example of a sticky organization. When the IMF was initially created, the IMF's main purpose was to manage the international monetary system referred to as the Bretton Woods system after World War II. Yet the Bretton Woods system lasted for only 25 years and collapsed in 1971 when the IMF lost its initial purpose to exist. However, instead of ceasing its operation, the IMF managed to successfully transform itself into a loan agency with growing emphasis on developmental assistance and new kinds of lending facilities. Third, applying the principal-agent theory, Keohane and Martin argue that there is room to maneuver for international institutions, independent from the sovereign states that once created them. The argument is fully expanded in the delegation and agency literature introduced in the third section.

Realism

Studies in the realist tradition start with the same premises that states act rationally and the international system is anarchic and characterized by the prisoner's

dilemma, but realist studies arrive at a very different conclusion. At its extreme, neorealism does not acknowledge the roles that international institutions play in international relations as realism emphasizes the quintessential role of state power in world politics. Thus, for some neorealists, international institutions exert little independent influence; they just represent the most powerful state's interest and are only effective when the most powerful state allows them to be so. For others, institutions have some effects over international affairs, but the primary determinant of the outcome is still the power of states.

Realists, often associated with hegemonic stability theory, argue that international organizations are created by the most powerful states in international system to function as instruments for furthering powerful states' own interests. For instance, they would claim that the United Nations and the Bretton Woods system, including the IMF and the General Agreement on Tariff and Trade, were created by the United States to provide stability in security, trade, and financial affairs. With these organizations, the United States could reap disproportional benefits and consolidate its superpower status. Moreover, these international organizations may have been used as the American foreign policy instruments.

Gilpin (1981) argues that a hegemon creates international institutions to maintain the global order by providing public goods, most prominently free trade and stability. But providing public goods is not costless, and this cost is generally associated with the slow decline of the hegemon. As the hegemon starts to decline, simultaneously, another secondary power or powers start to rise and threaten the status of the hegemon. According to him, this is the temporal dynamic of the rise and fall of the hegemon.

Gruber (2000, 2001) also emphasizes how power is the driving force in creating an international organization. He argues that more powerful states can create an international institution without consents from other states and then can force other states to join the international institution by making them worse off when they resist participating. This can be done because powerful states have go-it-alone power, moving the status quo point to a new point where other states, maintaining the status quo ante, obtain a worse payoff. Gruber (2001) explains the initiation of the North American Free Trade Agreement (NAFTA) in this way. In between relatively weaker countries—Canada and Mexico—the United States managed to manipulate and limit choices within the choice sets for Canada and Mexico. That is, while Canada and Mexico wanted to maintain the status quo ante (that is, no free trade agreement), the United States managed to make it so that there were only two available options for each state: being left out by the bilateral free trade agreement between the United States and the other state or participating in the three-way free trade agreement. For instance, when the United States made the bilateral free trade agreement with Canada, Mexico was left with the option of joining the three-way free trade agreement or being left out of its trade share in the U.S. market.

Between these two choices, each state decided to choose the less harmful option, which is participating in the three-way free trade agreement and creating NAFTA.

Other realists argue that although international organizations may influence international affairs to a certain extent, the form of an international organization is heavily determined by state power. Targeting the efficiency argument proclaimed by Keohane and other neoliberal institutionalists, Krasner (1991) points out that capturing the international bargaining process with the prisoner's dilemma is misleading and that the inherent problem in international cooperation is seldom to reach an international agreement but more often to decide what kind of an agreement is reached. Then he argues that powerful states often dictate the rules of interactions, such as who relevant players are and who proposes an agenda first when bargaining an international agreement. In this way, powerful states dominate and reap the most gains when cooperation occurs. He illustrates the point with the case of the distribution of global communication resources. He demonstrates how powerful states often monopolized the use of radio waves, space for satellites, and magnetic waves when weaker states had no technology to demand more equal distributions of those limited resources. Only after weaker states developed technologies to intervene the monopolistic use of these limited international goods by powerful states did the powerful states concede.

With regard to the effectiveness of international organizations, Mearsheimer (1994–1995) argues that international institutions are epiphenomenal at best. Mearsheimer (1990) proclaims that international institutions are just reflections of the most powerful state's interest, and once its interest is evaporated, there is no use for the international institutions anymore. For instance, he predicts the instability of Europe after the collapse of the Soviet Union, on the grounds that the North Atlantic Treaty Organization would not function properly any longer because the most powerful state, the United States, had lost its interest to maintain stability in Europe when the Soviet Union had collapsed.

Constructivism

Constructivists take a very different approach to provide powerful critiques of both neoliberal institutionalism and realism. First of all, constructivists contend that the preferences of states are not fixed, as neoliberal institutionalists and realists often assume, but are socially constructed through repeated interactions among states. Similarly, the rules of the interactions among states are shaped and transformed through states' interactions. For instance, Wendt (1999) argues that the logic of anarchy is not fixed but varies and is constituted by interactions among actors within the international system, and the preferences of states and the prospect of cooperation vary accordingly. Specifically, under the Hobbesian culture of anarchy, states perceive each other as an enemy; thus, international cooperation among states would not be likely. In comparison, states perceive

each other as rivals under the Lockean culture of anarchy; thus, cooperation can be achieved with additional institutional settings such as international organizations. Finally, under the Kantian culture of anarchy, states see each other as friends, and states' interests are in harmony with each other, making it unnecessary to collaborate. In general, studies of international organizations in constructivist tradition emphasize the roles that nonmaterial factors, such as identity, norms, cultures, and ideas, shape and transform states' interests, how the transformed states' interests affect international organizations, and how international organizations transform the states' interests in turn.

Haas (1989) illustrates how the spread of scientific knowledge by the epistemic community¹ has altered states' interests over time and how it allowed states to create the Mediterranean Action Plan, the international arrangement among southern European and North African countries to regulate international marine pollution in the Mediterranean Sea. He shows that power- and interest-based theories cannot fully account for the adaptation and implementation of the Med Plan. Then, he proposes the alternative argument that the United Nations Environmental Programme assisted creation and empowerment of the ecological epistemic community and shows how the empowered epistemic community was able to spread scientific knowledge across the Mediterranean countries and persuade these governments. This allowed the change of preferences of those countries that initially had opposed the Med Plan, such as Algeria and Egypt, and led them to ultimately participate in and comply with the Med Plan. A similar argument is made by Finnemore (1993). She argues that international organizations often have their own goals and set and spread norms accordingly. In the empirical study, she shows that the United Nations Educational, Scientific, and Cultural Organization (UNESCO) was able to spread the norm that modern states needed to take science seriously. As more and more states accepted the norm, UNESCO was able to persuade many developing countries to establish an independent scientific bureaucracy.

Johnston (2001) highlights a role of an international organization play that is not emphasized by neoliberal institutionalists or realists. He asks how involvement in international institutions changes states' behavior in the absence of material rewards and punishment in his study and answers the question with what he calls *socialization process*. Johnston applies Siegal's definition of socialization, "the process by which people learn to adopt the norms, values, attitudes, and behaviors accepted and practiced by the ongoing system" (as cited on p. 495), to international relations and argues that through socialization, states can change their preferences, which may lead to more cooperative behaviors. Through providing the forum for socialization and facilitating persuasion and social influence, such as social opprobrium and back patting, international institutions play the important role in helping to promote cooperation in a way that is often neglected by neoliberal institutionalists.

Bringing Theories Together

This section lists theories of international organizations that propose interest-, power-, and knowledge-based explanations—neoliberal institutionalism, realism, and constructivism, respectively—of creation and functions of international organizations. On one hand, power-politics arguments discount the roles that international institutions may play in world politics. In comparison, interest-based arguments emphasize how international institutions provide information so that participating states can achieve mutually beneficial cooperation. More recently, constructivists provide alternative arguments that are often complementary to the existing rationalist arguments.

Although these diverse arguments have advanced our understanding of many aspects of international organizations, they are often presented as competing theories of international organizations. Is it, then, impossible to bring these diverse arguments together in order to achieve more comprehensive understanding of international institutions? Hasenclever, Mayer, and Rittberger (2000) propose one possible way to bring the theories together that they term as *contextualized rationalist theory*. They see the core difference between realist and neoliberal institutionalist arguments in terms of how much states care about how much they gain from cooperation in absolute terms as opposed to how much states gain relative to other states. Since states are likely to behave as neoliberal institutionalists argue when they are motivated by gains in absolute terms and, conversely, states are less likely to cooperate when they are motivated by relative gains, adding a theory of states' motivations prior to considering the two rationalist explanations may allow scholars to apply an appropriate theory to understand a given research question. Furthermore, studies in the constructivist tradition may be able to provide useful theories of states' motivation that realists and neoliberal institutionalists make assumptions about and treat as given. This may be one way to bring constructivism and the rationalist theories together in a constructive manner.

Rational Design and Delegation

Rational Design Literature

Building on the earlier work on the roles of international institutions in facilitating cooperation among states and moving forward from the once-heated debates regarding whether international institutions matter, the more recent scholarship in the study of international organizations shifts its focus to variations in features of diverse international organizations.

In the 2001 *International Organization* special issue, Koremenos, Lipson, and Snidal (2001a) laid out the platform of rational design of international institutions. For these scholars, not only the very existence of international

institutions is intended by sovereign states, but also their specific characteristics, whether membership, issue scope, (de)centralization, decision-making procedures, or flexibility, are also rationally designed by participating states. This argument is based on the observation that institutions are outcomes of many rounds of deliberate negotiations among engaged states.

There are indeed enormous variations in characteristics of international organizations. For instance, some international organizations such as the European Union or NAFTA are regional, while other international organizations such as the WTO or the United Nations are global. Some international organizations grant each state an equal vote, while others have some form of a weighted voting system: Some require simple majority, some require supermajority, and the others require unanimity. Some institutions have relatively strong centralized authority and some enforcement mechanisms, while other institutions provide only nonbinding consultations.

For example, Rosendorff and Milner (2001) argue that the escape clause of the WTO procedure is deliberately chosen by member states to cope with domestic uncertainty on preference changes. Without the escape clause, a potential domestic preference shift among different social groups may force their governments to withdraw from the WTO even if the overall cost-benefit calculation of continuing participation in the regime is still bigger than that of quitting the WTO. Individual states are uncertain about future domestic preference dynamics when they sign the agreement. With the risk of abrupt quitting of trading partners from the trade regime, states would be hesitant to sign the agreement in the first place. Facing this dilemma, those who negotiated and designed the WTO created the escape clause to facilitate temporary needs of member states to violate the agreed-on rules. But this escaping from the established rules should not be costless, since if it were, then states could resort to the escape clause more often than necessary, and consequently, this will reduce overall benefits for all participating states. Thus, the existence of the costly escape clause in the trade agreement is the optimal solution for the uncertain future of the domestic preference dynamics. The escape clause is created in such a way to accommodate states' needs to accommodate temporary domestic pressure while maintaining the overall agreement structure.

In a similar vein, the North American Treaty Organization (NATO) expansion study by Kydd (2001) highlights how the restrictive NATO membership criteria of democratization work as the filter that enable potential members to signal their strong commitment to security cooperation and keep out problematic members who would be less cooperative in NATO's operations. The membership criteria of NATO serve as the screening mechanism that allows only those countries genuinely committed to join the membership. One can argue that a similar mechanism exists for the European Union. To join the European Union, candidate states need to demonstrate that they are committed to

noninflationary economic policies, respect for human rights, and democracy, among others.

Delegation and Agency

Another line of emerging literature in the studies of international organizations looks at the tension between states that create and finance international organizations and the international organizations within the principal agent framework. States are the principals, and they decide to delegate some of their authority to the agents (international organizations) because it is often more efficient to have agents with better expertise in a particular issue area handle everyday operations.

The issue of delegation holds significant practical implications. On one hand, pundits and academics blame international organizations for not being very effective or for being too biased. And the major reason of the ineffectiveness and bias is states' tight control over organizations. Often, international organizations are under tight control of sovereign states, and unless states agree to allow international organizations to act, there is little difference that international organizations can make. For instance, peacekeeping missions of the United Nations are often proven ineffective, with too-small and often too-late engagements in humanitarian crisis situations. And this ineffectiveness comes from the fact that peacekeeping missions are possible only when states agree to make them and are willing to send their own civil and military personnel to conflict-ridden regions. In a different context, the IMF is often accused of being biased toward its major shareholders, such as the United States and other G-7 countries since the IMF's decision making is largely controlled by these states.

On the other hand, others accuse international organizations of being too independent and too irresponsible once they are created. International organizations are depicted as pursuing their own interests, whether the interests are ideological or bureaucratic. International organizations are funded by citizens' taxes through member states' contributions, yet international organizations are too autonomous and pursue their own goals rather than serving those who finance them. Furthermore, international bureaucrats are not electorally responsible yet often make public policy decisions that exert enormous influence over the welfare of ordinary citizens. For instance, the IMF often consults borrowing countries' governments to formulate IMF programs that include policy reform measures such as privatization, civil service and public sector reforms, trade liberalization, pension reforms, and labor market reforms. Thus, the IMF staff often exerts enormous influence over policy decisions yet is not accountable to those who are affected. In sum, international organizations are criticized on two contradictory grounds—one, international organizations are under too-tight control of strong states and thus are not very effective, and two, international organizations are too autonomous from the states and have little accountability.

To better understand the relationship between states and international organizations as principals and agents, scholars ask questions like the following: Why do states delegate certain tasks and responsibilities to international organizations rather than acting unilaterally or cooperating directly without international organizations, and how do states control international organizations once authority has been delegated to the organizations? Studies in the edited volume *Delegation and Agency in International Organizations* (Hawkins, Lake, Nielson, & Tierney, 2006) are a culmination of such recent scholarly efforts. Through careful examinations of diverse international organizations such as the IMF, the European Union, the United Nations, the World Health Organization, and the WTO, the collective efforts in the volume conclude that successful delegation to international organizations is possible, given that states have and maintain sufficient information and that little conflicting interests between international organizations and states exists.

Conclusion

This chapter introduces the definitions of essential terms in studies of international organizations and then briefly discusses existing theories of international organizations. There is an increasing number of studies that have informed us about why states create international organizations, what influences the features of international organizations, what functions international organizations perform once they are created, and how states oversee or how independent international organizations are from states' control.

Although we know a lot more about international organizations, there are also limitations of the current scholarship and hence many fruitful areas that future scholarship can address. First of all, a more comprehensive understanding of politics of international organizations may be possible when we study the interaction between the formation of state preferences and the creation and functions of international organizations. Proper attention to domestic politics can prove useful to understanding formation of state preferences. Alternatively, international norm dynamics and evolution of identity of interacting states can be fruitful to explain state preference formation and transformation. Second, the current scholarship of international organizations is state-centric and pays little attention to nonstate actors. There are a very limited number of studies that investigate nongovernmental international organizations, even with the growing importance of nongovernmental international organizations in world politics. Thus, more attention to the partnerships between formal international organizations and nongovernmental organizations is needed. And studies that investigate the conditions under which those partnerships may be proven fruitful will have huge policy implications for the future cooperation between nongovernmental international organizations and formal international organizations.

Studies of international organizations have made major advances for the past three decades, but there are still many exciting yet unresolved questions to be explored. Moreover, as the roles that international organizations play in international relations keep growing and issue areas that the international organizations participate in increase, there will be even more need for proper understanding of these organizations. Thus, the subfield of international organization will be a very exciting place to develop research programs for current and future political scientists.

Note

1. The term *epistemic community* is defined as a network of knowledge based experts or groups with an authoritative claim to policy relevant knowledge within the domain of their expertise. Members hold a common set of causal beliefs and share notions of validity based on internally defined criteria for evaluation, common policy projects, and shared normative commitments (Haas, 1989).

References and Further Readings

- Abbott, K. W., & Snidal, D. (1998). Why states act through formal international organizations. *Journal of Conflict Resolution*, 42, 3–32.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Barnett, M. N., & Finnemore, M. (2004). *Rules for the world: International organizations in global politics*. Ithaca, NY: Cornell University Press.
- Finnemore, M. (1993). International organizations as teachers of norms: The United Nations educational, scientific, and cultural organization and science policy. *International Organization*, 47, 565–597.
- Gilpin, R. (1981). *War and change in world politics*. Cambridge, UK: Cambridge University Press.
- Gruber, L. (2000). *Ruling the world: Power politics and the rise of supranational institutions*. Princeton, NJ: Princeton University Press.
- Gruber, L. (2001). Power politics and the free trade bandwagon. *Comparative Political Studies*, 34, 703–741.
- Haas, P. M. (1989). Do regimes matter? Epistemic communities and Mediterranean pollution control. *International Organization*, 43, 377–403.
- Hasenclever, A., Mayer, M., & Rittberger, V. (1997). *Theories of international regimes*. Cambridge, UK: Cambridge University Press.
- Hasenclever, A., Mayer, M., & Rittberger, V. (2000). Integrating theories of international regimes. *Review of International Studies*, 26, 3–33.
- Hawkins, D. G., Lake, D. A., Nielson, D. L., & Tierney, M. J. (Eds.). (2006). *Delegation and agency in international organizations: Political economy of institutions and decisions*. Cambridge, UK: Cambridge University Press.
- Johnston, A. I. (2001). Treating international institutions as social environments. *International Studies Quarterly*, 45, 487–515.
- Karns, M. P., & Mingst, K. A. (2004). *International organizations: The politics and processes of global governance*. Boulder, CO: Lynne Rienner.
- Karns, M. P., & Mingst, K. A. (2006). *The United Nations in the 21st century*. Boulder, CO: Westview Press.
- Keohane, R. O. (1984). *After hegemony*. Princeton, NJ: Princeton University Press.
- Keohane, R. O. (1988). International institutions: Two approaches. *International Studies Quarterly*, 32, 379–396.
- Keohane, R. O., & Martin, L. L. (2003). Institutional theory as a research program. In C. Elman & M. F. Elman (Eds.), *Progress in international relations theory: Appraising the field* (pp. 71–108). Cambridge: MIT Press.
- Koremenos, B., Lipson, C., & Snidal, D. (2001a). The rational design of international institutions. *International Organization*, 55, 829–857.
- Koremenos, B., Lipson, C., & Snidal, D. (Eds.). (2001b). *The rational design of international institutions*. Cambridge, UK: Cambridge University Press.
- Krasner, S. D. (Ed.). (1983). *International regimes*. Ithaca, NY: Cornell University Press.
- Krasner, S. D. (1991). Global communications and national power: Life on the Pareto frontier. *World Politics*, 43, 336–356.
- Kydd, A. H. (2001). Trust building, trust breaking: The dilemma of NATO enlargement. *International Organization*, 55, 801–828.
- Martin, L. L., & Simmons, B. A. (2001). *International institutions: An international organization reader*. Cambridge: MIT Press.
- Mearsheimer, J. J. (1990). Back to the future: Instability in Europe after the cold war. *International Security*, 15, 5–56.
- Mearsheimer, J. J. (1994–1995). The false promise of international institutions. *International Security*, 19, 5–49.
- Milgrom, P. R., North, D. C., & Weingast, B. R. (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Economics and Politics*, 2, 1–23.
- Oye, K. (1986). Explaining cooperation under anarchy: Hypotheses and strategies. In K. Oye (Ed.), *Cooperation under anarchy* (pp. 1–24). Princeton, NJ: Princeton University Press.
- Rosendorff, B. P., & Milner, H. V. (2001). The optimal design of international trade institutions: Uncertainty and escape. *International Organization*, 55, 1019–1049.
- Ruggie, J. G. (1998). What makes the world hang together? Neo utilitarianism and the social constructivist challenge. *International Organization*, 52, 855–885.
- Stein, A. (1982). Coordination and collaboration: Regimes in an anarchic world. *International Organization*, 36, 299–324.
- Thompson, A. (2009). *Channels of power: The UN Security Council and U.S. statecraft in Iraq*. Ithaca, NY: Cornell University Press.
- Union of International Associations. (2007). *Yearbook of international organizations*. Available at <http://www.uia.be/yearbook>
- Wendt, A. (1999). *Social theory of international politics*. Cambridge, UK: Cambridge University Press.

52

INTERNATIONAL LAW

TINA KEMPIN REUTER

Christopher Newport University

International law is the body of legal rules, regulations, standards, and principles that govern international relations between or among states and other international actors. It deals with important concepts such as sovereignty (supreme authority over a territory); agreements and disputes between international actors; the use of force and self-defense; the regulation of the high seas, air, and space; international trade; and human rights. The United Nations (UN) and other international organizations have created a network of instruments addressing most aspects of international relations. International law influences large parts of everyday life—it makes it possible for us to send a letter to someone on the other side of the world, to travel internationally by just using our passports, and to know what time it is anywhere on the planet (for more ideas, see American Society of International Law, 2006).

International law is also known as *public* international law, which is distinct from private international law. Private international law deals with cases within the domestic legal systems of states, in which foreign elements are involved. Private international law addresses private matters, such as business disputes and family law, across international borders. Courts have to determine what jurisdiction and which laws apply to a certain case and how to enforce foreign judgments. By contrast, public international law, which is the focus of this chapter, is a separate legal system with its own rules and processes.

Public international law is in many respects comparable to domestic law, but it has just as many differences. Most importantly, international law is horizontal and decentralized. International law is horizontal because the main actors—namely, states—are considered equal in rights and duties and not subordinate to a higher authority (with some exceptions, e.g., the European Union [EU] member states). The system is decentralized because there is not one single executive, legislative, and judicial structure responsible for the creation, interpretation, and enforcement of international law but rather, as is shown in the following sections, different bodies and different processes (see Brownlie, 2003; Cassese, 2005; Evans, 2006; Shaw, 2003).

Nature and Development of International Law

International Law and International Relations

International lawyers and international relations scholars often study the same issues. However, they have traditionally operated in different spheres and did not profit from each other's work. Only in the past years has a growing consensus developed that collaboration and interdisciplinarity is fruitful and beneficial to both disciplines (see, for example, Anne-Marie Slaughter's work). In fact, both disciplines are divided and united by the same theoretical

debates. The core paradigms—realism, liberalism, and constructivism—are mainly the same with regards to the basic assumptions about the world (see Chapter 36 in this volume, “History of International Relations,” for an overview of international relations theory). Realists generally have a critical view of international law because they believe that power and state capabilities (particularly military strength), rather than common interests, shape international relations. Cooperation occurs only if it is in the interest of states and has nothing to do with universal values or principles shared by all states in the world. As a result, the likelihood of cooperation is ever changing, depending on the status of international politics, the balance of power, and the relative strength or weakness of a given state. Liberals, by contrast, see the development of international law and international institutions as a result of increasing economic and social interdependence of states. Shared values such as peace, human dignity, and liberal democracy are important factors contributing to collaboration among state and nonstate actors. International organizations, nongovernmental organizations (NGOs), and individuals play an important role because they often serve as agents representing these fundamental values (e.g., human rights or environmental protection). Finally, constructivists see the international system as a construct of international actors that share the same ideas, conceptions, and values. Legal norms play an important role shaping the identity and interests of actors, which in turn influence the structure, institutions, and processes of the international system. International actors define themselves through social interactions, which change as values change (for more details, see Armstrong, Farrell, & Lambert, 2007).

International Law and National Law

As mentioned previously, international law is an independent legal system. Nevertheless, it interacts with domestic law in more than one way. Much of international law has to be incorporated into national law and depends on the enforcement by domestic legal systems. There are two ways to integrate international law into national law: monism and dualism. Monism sees international law and domestic law as parts of the same legal system, in which international law trumps national law. The closest approach to a monist tradition is used in the Netherlands, where no transformation into national law is needed and international law can directly be invoked by the citizens or applied by a judge. Dualism holds that international law and domestic law are separate legal systems, which operate on different levels and are based on different premises and processes. Consequently, all international norms have to be transferred into national law. In the United Kingdom, for example, international law is only applicable in British law once an act of Parliament is passed to give effect to it.

The United States’ position is generally dualist but not as pronounced as in the United Kingdom. Domestic law is considered higher than international law, which means that

U.S. law could be in violation of its international treaty obligations and, if so, leads to international consequences such as condemnation, protest, and sanctions. However, Congress enacts laws generally in concordance with international obligations, and breaches tend to be very rare and minor in scope. International treaties are recognized as “supreme Law of the Land” (Article VI, Section 2, of the U.S. Constitution). Treaties are negotiated by the president but can be ratified only with the approval of two thirds of the Senate (Article II), except in cases of executive agreements, which the president can negotiate within his own authority (see the following for a more detailed description of the ratification process). A treaty can be self-executing or non-self-executing. Self-executing treaties are directly part of the law of the land and do not require domestic legislation for the treaty to enter into force. Non-self-executing treaties require an act by Congress. Whether a treaty is self-executing or not is determined by the intentions of the signatories and the interpretation of the courts. As a general rule, the more detailed and specific a treaty, the higher the probability that it is self-executing. Parts of the UN Charter (1945), for example, have been deemed non-self-executing because the rules are too broad and ambiguous to be directly applied in U.S. courts (for more details, see Bederman, 2006).

History

Modern international law has its roots in the Peace Treaty of Westphalia (1648). The Peace of Westphalia brought an end to European religious wars (it concluded the Thirty Years’ War) and coincided with the rise of nation-states. The rulers decided that each entity should be sovereign, which meant that each sovereign (or in democracies today, the people) enjoyed supreme authority over his or her territory. This became the key concept of the international system: States are independent and autonomous and, because of sovereignty, equal in their ability to enter international relations.

Relations and agreements among sovereign entities can be traced back to ancient times, particularly regarding issues relating to peace and war, trade relations, and the law of the sea (including the prohibition of piracy) (for an overview, see Grewe, 2000). The Greeks heavily influenced international law by introducing the concept of natural law: the notion that some laws are given by nature and therefore are fixed and universal. According to the natural law tradition, what is right or wrong remains the same and does not vary over time or if the political or cultural context changes. This concept inspired modern human rights law. Roman law laid the basis for *jus gentium* (law of nations), which regulated the status of foreigners and relations between foreigners and Romans. During the Middle Ages, natural law, inspired by religious notions, became the intellectual foundation of the law of nations.

With the beginning of the rise of nation-states, scholars in Europe sought to define the content and existence of international law (see also Murphy, 2006). The first to

discuss international law were the Spanish Scholastics such as Francisco de Vitoria (1486–1546) and Francisco Suarez (1548–1617), who relied heavily on Catholic theology in their interpretations of international law (which failed to convince the Protestant community). The Italian legal scholar Alberico Gentili (1552–1608) is considered the father of the secular school of thought in international law, initiating a transformation from a theological interpretation to a concept of secular philosophy based on reason. Following his footsteps, the so-called father of international law, Hugo Grotius (1583–1645), established international law as a comprehensive system in his 1625 masterpiece, *De Jure Belli ac Pacis* (*On the Law of War and Peace*). Many scholars followed Grotius's natural law argumentation, most prominently the German jurist Samuel von Pufendorf (1632–1694), who insisted on the natural equality of states. German philosopher Christian Wolff (1679–1754) and Swiss legal scholar Emerich de Vattel (1714–1767) both paid tribute to the natural law tradition but characterize law just as much as the result of state practice. This new understanding of law, known as *positivism*, shaped international law in the 18th and 19th centuries.

Positivism states that law comprises the rules made by human beings (in the case of international law, by states) and stresses the importance of sovereign consent to international norms. Positivists argue that law is not given by nature or connected to ethical and moral beliefs but consists of the body of rules to which states have agreed (a treaty, for example, is proof of that consent). The influence of positivism in international law, however, was undermined by the two world wars, the creation of the two major international organizations—the League of Nations in 1919, after World War I, and the United Nations in 1945, after World War II—and the rise of human rights (the revival of the natural law tradition). International law became truly global in the years after the founding of the UN, when decolonization ended in the establishment of independent states in areas formerly controlled by European powers. The ideological conflicts (liberal democracy and free market economy versus communism and centrally planned economy) during the cold war and the cultural influence of the newly independent nations in Africa and Asia influenced the development international law and extended it beyond its mainly European origin and background. International law today addresses issues such as decolonization, racial discrimination, the rights and responsibilities of individuals, and economic aid.

After the collapse of the Soviet Union, many predicted a new era for international law. Two trends became particularly significant. First, the importance of human rights was strengthened by international interventions in reaction to humanitarian disasters (e.g., those in Kosovo in 1999, Rwanda in 1994, and Somalia in 1992). Second, the increasing interconnectedness among international actors due to globalization led to new areas of international cooperation such as the struggle against climate change or the

improvement of trade relations. The most important issues in international law at the beginning of the 21st century are human rights and the fight against human rights violations (international criminal law and international humanitarian law), dealing with environmental problems, and the ongoing discussions on the use of force.

Sources

The sources of international law are laid down in Article 38 of the Statute of the International Court of Justice (ICJ, 1945). These sources are international conventions (treaties), international custom, general principles, and judicial decisions and teachings of highly respected scholars in the field. The latter (decisions and teachings) are secondary sources of international law, while treaties, custom, and general principles are considered primary sources (for more information generally, see Bederman, 2006; Brownlie, 2003; Cassese, 2005; Lowe, 2007; Murphy, 2006; Shaw, 2003). The following paragraphs focus on primary sources of international law.

Treaties

Treaties are written agreements between states. The signatories consent to be bound by its content, which makes a treaty legally binding. Treaties have many different names, including convention, agreement, pact, accord, protocol, charter, and covenant. Depending on the number of signatories, international law distinguishes between bilateral (two states) and multilateral (three or more states) treaties. Multilateral treaties, such as the UN Charter, the Geneva Conventions, the Nuclear Non-Proliferation Treaty (NPT), the Convention on the Prevention and Punishment of the Crime of Genocide (Genocide Convention), the Rome Statute establishing the International Criminal Court (ICC), and the UN Convention on the Law of the Sea (UNCLOS), to name a few, usually have higher significance than bilateral treaties (with some exceptions, e.g., the disarmament agreements between the United States and Russia). The purpose and scope of treaties can be broad (e.g., the UN Charter) or specific (e.g., the International Convention for the Regulation of Whaling).

Treaty making is a multitiered process. After the text of the treaty has been negotiated, a treaty has to be signed and ratified before it enters into force. Treaties are usually signed by the head of state or his or her representatives. Ratification procedures vary depending on the constitutional structure of the state, but in most cases, the parliament or the people have to express their consent (see the aforementioned procedure in the United States). Multilateral treaties often include special clauses specifying the process (e.g., a treaty enters into force after a given time or after a set number of ratifications has been reached). Once a treaty is signed and ratified, the state has the obligation

to execute it in good faith (a principle known as *pacta sunt servanda*, Latin for “agreements must be kept”).

Some treaties allow states to withdraw from its obligations (e.g., Article X of the NPT the so-called exit clause) or to opt out of a particular provision (termed *reservations*). Treaties can also be terminated or suspended by consent of the parties, which is more difficult in the case of multilateral treaties. If the circumstances of a treaty change dramatically, a party to a treaty can withdraw or terminate the treaty unilaterally (a principle called *clausula rebus sic stantibus*, Latin for “things standing thus”). However, if the fundamental change of circumstances is a result of one party violating the treaty, the principle does not apply.

Three approaches are used to understand and interpret the meaning of a treaty. First, the textual approach interprets the ordinary meaning of words of the provision as expressed in the text of the treaty and as derived from the context. It also takes different translations of the same treaty into account (e.g., UN treaties are legally binding in all official UN languages). Second, the intentional approach relies on the *travaux préparatoires* (French for “preparatory works”) and focuses on the intent of the drafters when they crafted the treaty. And third, the teleological approach seeks to understand the purpose of the treaty, meaning that a treaty provision should be interpreted in the way that makes the treaty most effective. Generally, all three approaches are used in combination.

Customary Law

The ICJ (1945) Statute refers to custom “as evidence of general practice accepted as law” (Article 38). To prove that a rule is customary, courts have to consider the actual practice of states and the acceptance by states of the practice as law (termed *opinio juris sive necessitates*, Latin for “opinion of law or necessity”). State practice is established by the actual behavior of states—namely, whether states use a practice repetitively and consistently over time. *Opinio juris* requires that states follow a rule because they believe it is lawful to do so, and not only because it is in their self-interest or because they are pressured to do so. To prove *opinio juris* is very difficult because states usually do not explicitly state why they follow a rule and because reasons for adherence overlap. Although all states contribute to custom, it is the practice of the most powerful states that determines the main course of the development of international customary law. Once a rule becomes customary, all states are bound by it, regardless if they expressed their consent (e.g., ratified a treaty). Most customary rules are supported by widely accepted treaty provisions (e.g., common Article 3 of the Geneva Conventions, discussed in a following section).

The concept of *jus cogens* (Latin for “compelling law”) is part of customary law. *Jus cogens* are peremptory norms that have to be followed by states in any situation. They enjoy a higher status than normal customary rules. The specific scope of *jus cogens* rules is unclear, but international

tribunals have indicated that *jus cogens* includes the prohibition of slavery, genocide, forced disappearances, piracy, aggression, and state-sponsored torture. Related to *jus cogens* is the concept of rules *erga omnes* (Latin for “toward all”). Breaches of *erga omnes* rules concern all states and can be prosecuted anywhere by any state, regardless of where and by whom the crime was committed.

General Principles

The third primary source of international law is general principles. General principles provide a mechanism to address issues that are not already regulated by treaties or by customary international law. Many general principles arise from comparable practices in domestic law and concern aspects of the judicial process. The most important principle is that of good faith—namely, the fact that states have to execute their treaty obligations honestly and to the best of their ability.

Legally Nonbinding Rules (Soft Law)

The traditional sources of international law do not account for legally nonbinding norms. However, so-called soft law can have an important impact on international relations and on the development of international law. The Universal Declaration of Human Rights (UDHR), for example, is a legally nonbinding document but has had a tremendous influence on the development of international human rights law. Legally nonbinding norms can be found in treaty provisions that are too unspecific to be implemented automatically (e.g., preambles), political pacts (e.g., the Final Act of the 1975 Conference on Security and Cooperation in Europe [CSCE] in Helsinki), resolutions of the UN General Assembly (but UN Security Council resolutions are legally binding), and codes of behavior for states and nonstate actors (e.g., food safety codes or codes of conduct for multinational enterprises). The advantage of soft law over hard law is that states are often more inclined to negotiate and conclude such documents. Legally nonbinding norms have no special procedures or bureaucratic obstacles attached to them (as opposed to treaties; see previous discussion), and states can be less cautious about the details of the agreement. Despite being legally nonbinding, such rules are generally obeyed (they are seen as politically binding).

Subjects of International Law

Subjects of international law are the actors of international relations that bear international legal rights and responsibilities. By contrast, objects of international law are the entities, processes, and instruments acted on (the topics of international law administration). Traditionally, states were the only recognized subjects of international law, but over time, other actors such as international organizations,

individuals, and other nonstate actors have gained international legal personality (see, generally, Bederman, 2006; Brownlie, 2003; Cassese, 2005; Lowe, 2007; Murphy, 2006; Shaw, 2003). Today, the line between subjects and objects of international law is blurred (individuals, for example, are both objects and subjects of international law).

States

States are the primary subjects of international law and accordingly possess the most extensive obligations and rights in the international system. To qualify as a state, an entity has to meet four criteria (as outlined in the 1933 Montevideo Convention): a permanent population, a defined territory, a government, and the capacity to enter into relations with other states. These principles are still valid today, although a state does not automatically lose statehood in case of border disputes or major migration (e.g., refugee flows).

Population

The size or density of a population does not matter when determining statehood. The only criteria are the relative permanence (this includes nomadic peoples) of the population and the requirement that the people form a living community. The staff of the Roman Catholic Church in Vatican City, for example, qualifies, which is one of the reasons why the Vatican is considered a state in international law. States award their citizens nationality, which individuals acquire in one of two ways: *jus sanguinis* (Latin for “right of blood”) or *jus soli* (Latin for “right of soil”). As the names indicate, in states using *jus sanguinis*, citizenship is inherited, given by blood from the parents to the children. The second approach bases citizenship on where a person is born. Typical emigration countries, such as states in Europe, prefer *jus sanguinis*, while traditional immigration states like the United States usually use *jus soli* to determine a person’s nationality.

Territory

The sovereign territory of a state includes airspace and, if the state is located on the sea, access to resources of the ocean. Internal waters such as harbors, lakes, and rivers and the sea up to 12 nautical miles from the baseline (the low-water mark on the coast) are under full sovereignty of the state. The territorial seas are subject to the right of innocent passage—namely, the right of ships to pass peacefully through these waters. Foreign ships in the territorial waters of a state may not threaten or use force, spy, deliberately pollute, violate immigration regulations, or use resources unauthorized (e.g., fishing). Beyond the territorial seas, the states can claim a contiguous zone, which extends another 12 nautical miles from the end of the territorial waters for the purpose of preventing or punishing

“infringement of its customs, fiscal, immigration or sanitary laws and regulations within its territory or territorial sea” (UN, 1982, Article 33, No. 1). The United States has claimed a contiguous zone in 1999. The exclusive economic zone extends for 200 nautical miles, within which a state has the right to access to all natural resources in the water and continental shelf (including fishing, mining, oil exploitation, scientific research, and artificial installations such as oil platforms). The state cannot control the access of ships, whether for peaceful or belligerent purposes, in the exclusive economic zone. If states are geographically close (adjacent or opposing, no international waters between them), the borders are determined by equidistance (line of equal distance from the baseline).

New states can be established through decolonization, split or dissolution (e.g., Czechoslovakia, Yugoslavia, and the Soviet Union), merger (North and South Yemen), absorption (of Eastern Germany into the Federal Republic of Germany), or annexation (e.g., Austria by Germany in 1938 or Tibet by China). The use of force to acquire territory is prohibited by the UN Charter (1945; see the following discussion). If states are created by the dissolution of existing states or in the case of colonial independence, the former administrative borders are the state’s new boundaries (a principle termed *uti possidetis*, Latin for “as you possess”).

Government and Capacity to Enter International Relations

A state’s government has to be effective, meaning that a state has to be able to fulfill the principle of self-determination and independence. Protectorates, self-governing territories (e.g., Puerto Rico), or states that have assigned their foreign relations to other states (e.g., the Principality of Liechtenstein to Switzerland) do not meet this criterion. The fact that self-determination plays such a crucial role defining states has led to the understanding that even some separatist movements enjoy recognition as actors of international law (e.g., the Palestine Liberation Organization).

Although not explicitly mentioned in the Montevideo qualifications, a state is evidently unable to enter international relations without recognition by other states. To enter international relations, states deploy ambassadors to serve as representatives of their interests. In some cases, state recognition is conditional on criteria such as democratic government, respect for human and minority rights, observance of existing borders, and commitment to disarmament and nuclear nonproliferation (see, for example, the criteria of the European Community in the Declaration on the Guidelines on the Recognition of New States in Eastern Europe and in the Soviet Union, 1991). Recognition of states has to be distinguished from the recognition of governments. The nonrecognition of a government usually leads to the discontinuation of diplomatic relations (e.g., the United States has no formal diplomatic relations with Iran).

Recognition of a state is a political, not a legal, decision. Kosovo, for example, which declared its independence in February 2008, has been recognized by many Western states but not by the traditional allies of Serbia. Abkhazia and South Ossetia, the two separatist regions of Georgia, have been recognized only by Russia and Nicaragua.

Nonstate Actors

Individuals

Individuals were only objects of classical international law. If the rights of individuals were violated by another state, the state could adopt claims on behalf of its nationals, a practice known as diplomatic protection. Since the establishment of the United Nations, individuals are subjects of international law, which means that they have rights and responsibilities. The rights are generally referred to as human rights; the responsibilities are covered by international criminal law.

Rights. Individual rights are laid down in various human rights documents and monitored by an array of institutions. The principal document is the Universal Declaration of Human Rights (UDHR), created in 1948, which establishes the basic principles and protections of individuals. The most important rights include the right to life, liberty, and equality; due process; freedom from discrimination and torture; freedom of speech, religion, and assembly; the right to self-determination of peoples; and minority rights. The rights in the UDHR were integrated in two treaties, the International Covenant on Civil and Political Rights (ICCPR), created in 1966, and the International Covenant on Economic, Social and Cultural Rights (ICESCR), created in 1966. In addition, specific human rights treaties focus on a particular crime (such as the Genocide Convention of 1948 and the Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment [CAT] of 1984) or on protecting a special group of people (e.g., the International Convention on the Elimination of All Forms of Racial Discrimination of 1965, the Convention on the Elimination of All Forms of Discrimination Against Women of 1979, and the Convention on the Rights of the Child of 1989). With the exception of the Genocide Convention, these treaties have monitoring bodies that oversee the implementation of the treaty. Other human rights bodies, such as the Human Rights Council (the Commission on Human Rights until 2006) or regional institutions implementing regional human rights treaties (the European Convention on Human Rights, a court; the Inter-American Convention on Human Rights, a commission and court; and the African Charter on Human and Peoples' Rights, a commission and court), protect and implement human rights (see, for more details, Claude & Weston, 2006; Donnelly, 2003; Steiner, Alston, & Goodman, 2008).

Responsibilities. Individual responsibilities generally relate to piracy and crimes implying individual criminal responsibility. Individuals can be held responsible for acts that constitute a violation of international obligations (e.g., the prohibition of hijacking, hostage taking, torture, or international terrorism) or so-called atrocity crimes, such as genocide, war crimes, crimes against humanity, and other offences covered by the statutes of the international criminal tribunals (the International Criminal Tribunal for the Former Yugoslavia [ICTY], the International Criminal Tribunal for Rwanda [ICTR], and the ICC). Individuals who commit any of these crimes are subject to universal jurisdiction, which means that perpetrators can be prosecuted for these crimes by any competent national or international court regardless of where the crime was committed and regardless of the perpetrator's nationality (see Cassese, 2008; Cryer, Friman, Robinson, & Wilmschurst, 2007).

International Organizations

International organizations are collectivities of states. They can take many forms—they can be bilateral, regional, or global in scope and might address narrow or broad concerns. Accordingly, the rights and duties of international organizations vary widely. While some international organizations enjoy international legal personality and are recognized as international actors (the UN, for example), others are not. To establish international legal personality, international organizations have to possess a constituent treaty (charter), are created on the basis of international law, and have an organ that is independent from its member states (see also Chapter 51, "International Organizations and Regimes," in this volume).

The first international organizations developed in the 19th century and regulated specific issues such as international mail and telecommunication (in fact, the Universal Postal Union was the first international organization, created in 1865), weight and measurement standardization, and prevention of diseases. The first universal organization that aspired global membership and covered a variety of international issues was the League of Nations (founded in 1919), the predecessor of the United Nations. The League of Nations made important progress with the codification of international law and contributed to the development of minority rights and rules addressing social services. The UN is today the most important international organization and has a major impact on the development of international law. The most important legislators are the UN General Assembly and the UN Security Council. The General Assembly enacts only legally nonbinding rules (see the aforementioned soft law), while all member states are bound by the decisions of the Security Council. The most important decisions of the Security Council are made in relation to Chapter VII of the UN Charter, which deals with threats to and breaches of peace or acts of aggression. The Security Council can authorize coercive (economic sanctions

or military intervention) and noncoercive measures (diplomatic interventions or good offices) to counter violations of international peace and security. In addition, regional organizations have added to the development of international law. Some organizations, most notably the European Union, have developed separate legal systems combining national and international law.

Special Topics

The Law of Armed Conflict

The law of armed conflict is generally divided into two categories: *jus ad bellum* (Latin for “just war”), which determines whether a state is allowed to use force, and *jus in bello* (Latin for “law of war”), which governs the law of armed conflict (see, generally, Kennedy, 2006; Kolb & Hyde, 2008).

Jus ad bellum: The Use of Force

Article 2 (para. 4) of the UN Charter (1945) prohibits the “threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the purposes of the United Nations” (Article 2, para. 4). What exactly constitutes the use of force is unclear, but according to a 1974 UN General Assembly resolution defining aggression, the use of force includes, among others, the invasion of one state of another (e.g., Iraq’s invasion of Kuwait), bombardment by one state of another (e.g., when NATO bombed Kosovo in 1999), blockade of another state’s coast or harbors (e.g., during the Cuban missile crisis in 1962), and attacks of another state’s armed forces. The only exceptions from the prohibition of the use of force are self-defense against other states (Article 51 of the UN Charter) or if the UN Security Council authorizes the use of force (under Chapter VII of the UN Charter). In the aftermath of the terrorist attacks on the United States on September 11, 2001, the UN Security Council extended the right to self-defense to international terrorism.

Beyond these two exceptions, the use of force is a controversial topic. Preemptive measures, for example, are only permitted if an attack is imminent. Or self-defense has to be necessary and proportional. What *imminent* and *proportional* mean depends on the circumstances, which gives room for various interpretations and misuse. The right to use force in humanitarian emergencies such as genocides or crimes against humanity without the authorization of the UN Security Council is subject of an ongoing debate. Recent developments point to the potential future permissibility of humanitarian interventions under certain preconditions and with UN involvement (see International Commission on Intervention and State Sovereignty, 2001). Regardless, if the use of force is legal

or not, states have to abide by the rules of the international law of war (see, for more details, Gray, 2008).

Jus in bello: International Humanitarian Law

International humanitarian law comprises the laws of armed conflict and defines the duties of states engaged in warfare, the obligations of neutral nations, and the rights, protections, and responsibilities of individuals. The most important documents are the Hague Conventions (1899 and 1907), regulating the conduct of warfare, and the Geneva Conventions (1949), focusing on the protection of victims of armed conflict—that is, wounded soldiers, prisoners of war, and civilians.

By the end of the 19th century, military technology advanced to a point where weapons could inflict unnecessary suffering on soldiers (e.g., so-called dum-dum bullets, which expand on impact and lead to extraordinary suffering and difficult medical treatment). In 1899 and 1907, heads of states met in The Hague to establish the rules on how states should conduct warfare (these meetings are generally referred to as Hague Peace Conferences). States laid down the principles of international humanitarian law—namely, the need to protect persons or property unless militarily necessary; the need to distinguish between military and nonmilitary (civilian) targets (known as *positive discrimination*); the need to identify actors under international humanitarian law (a group participating in warfare has to have a hierarchical structure, distinctive emblem, carry arms openly, and respect the laws of war); the rule of proportionality (military objectives have to be weighed against possible collateral damage to civilians and nonmilitary property); and in general the need for humanity in armed conflict. The Hague Regulations prohibit the use of poisonous gas; the killing of soldiers who surrendered; and attacks on undefended towns and buildings marked as religious, art, science, historic monuments, and hospitals. In addition, in case of occupation, the occupying force is responsible for the welfare of civilians living on the occupied territory and may not cease civilian property.

The four Geneva Conventions (1949) deal with the protection of the wounded and sick soldiers on land (I) and sea (II), prisoners of war (POW) (III), and civilians (IV). All four Geneva Conventions share common Article 3, which deals with “armed conflict not of an international character,” stating that POWs and civilians have to be treated humanely, may not be taken hostage or subjected to acts of violence, and have the right to due process. This is the only provision that applies to noninternational conflicts and is considered the minimum standard for warfare. An “armed conflict not of an international character” is any conflict that does not occur between two or more states. This includes internal conflicts, civil war, ethnic conflict, and international terrorism (e.g., al Qaeda versus the United States).

In addition to the Geneva Conventions, two Additional Protocols adopted in 1977 address the protection of victims of international armed conflict (Protocol I) and non-international armed conflict (Protocol II). The Additional Protocols revisit and clarify some of the Hague and Geneva regulations and expand the rules to new standards, which include, for example, the protection of the environment. The United States has not ratified either protocol on the grounds that they give too much protection and legitimacy to nonstate actors, particularly terrorists, but it has incorporated most of the Protocols' contents in its internal manuals. In 2005, a third additional protocol (Protocol III) was adopted to create a new emblem, known as the red crystal, to complement the already existing red cross and red crescent.

The implementation of international humanitarian law is the task of the International Committee of the Red Cross (ICRC, n.d.). The ICRC "is an impartial, neutral and independent organization whose exclusively humanitarian mission is to protect the lives and dignity of victims of armed conflict and other situations of violence and to provide them with assistance" (para. 1). Besides fostering the development of international humanitarian law, the ICRC plays an important role in monitoring prison standards, responding to humanitarian emergencies, and coordinating national Red Cross societies (see, for an overview of international humanitarian law, Dinstein, 2004, and the ICRC website).

International Economic Law

One of the most important objects of international law is the global economy. As mentioned previously, the establishment and regulation of trade relations between sovereign entities is one of the oldest areas of application of international law. International economic law includes the rules of international commerce, the regulations of the international trading and monetary systems, and the principles of international development and investment.

International commercial law derived from the *lex mercatoria* and the *lex maritima*, two bodies of law that existed since the Middle Ages. Today, international private law deals with international transactions, contracts, and dispute resolution between private actors across international borders.

International trade and monetary law is concerned with the global market. The primary focuses are tariffs (import taxes on foreign goods) and quotas (limitations of quantities of imported goods). The most important actors concerned with international trade and monetary law are the International Monetary Fund (IMF), the World Bank Group, and the World Trade Organization (WTO, the successor organization of the General Agreement of Tariffs and Trade, or GATT, regime). The IMF oversees the global financial system with the goal to stabilize international exchange rates and foster development. The World Bank

offers loans, advice, and support to developing states with the aim of reducing poverty. The key elements of the WTO are trade equality, dispute settlement, and the removal of tariffs. All three organizations have made an important impact on international relations, and some scholars see the future of international law in the regulation of the international economy (see, for example, Bederman, 2006; for a general overview, see Lowenfeld, 2008; Qureshi & Ziegler, 2007).

International Environmental Law

International environmental law has developed over the past decades because of the realization that a state should not use its territory or resources to harm another. Early cases deal with the regulation of air and river pollution, which led to the so-called polluter pays principle. The realization that many environmental issues cross borders and go beyond pollution led to a new awareness that what needed to be addressed was not only the consequences of human behavior on the environment, but also the values of environmental protection and guidelines for precautionary measures before action is taken. The central debate today is concerned with the relationship between economic development and environmental protection. Sustainable development—namely, to support economic growth in a way that is sustainable to the environment—is the attempt to reconcile the two seemingly opposing concepts.

International law has developed different approaches to environmental protection. First, liability and compensation regimes hold states responsible for actions that hurt the environment. These regimes have proven difficult to implement because they are very hard to negotiate. Second, the duty to consult and notify imposes the obligation on states to notify other nations about environmental problems and emergencies. Third, the regulation of pollutants, for example, by establishing emission limits, and the establishment of environmental standards that have to be met, has led to some progress. The 1997 Kyoto Protocol, part of the 1992 UN Framework Convention on Climate Change, set the goal of a 6% reduction of greenhouse gases from the 1990 levels by 2012. Fourth, the permission approach allows different scenarios depending on the pollutant (e.g., black lists of prohibited materials or white lists of materials with few or no restrictions; this scheme plays a role in regulating waste such as ocean dumping, chemicals, and pesticides). Along the same lines, the Kyoto Protocol established a scheme for emissions trade. Finally, the strongest approach is generally if environmental protection is linked to trade and economic interests. The Convention on International Trade in Endangered Species, for example, limited the trade of rare animals and plants and therefore eliminated incentives for investors (see, for an overview, Birnie, Boyle, & Redgwell, 2009).

Conclusion

International law has both widened and deepened since the beginning of modern international law in the 17th century. It has widened to include not only European nation states and European thought but also ideas, values, and issues from other areas of the world. International law today is truly global. At the same time, international law has deepened by going beyond the traditional issues of war, trade, and the law of the sea and includes now international environmental concerns and international criminal law. The sovereignty of states has become permeable—values such as human rights cross borders, and actors such as international organizations play increasingly important roles in international relations. New threats such as weapons of mass destruction, genocides, international terrorism, global poverty, and environmental issues will be a challenge for international law in the 21st century. These problems can be addressed only through a comprehensive approach, which will be made possible by the interconnection and collaboration of different parts of international law.

References and Further Readings

- Alvarez, J. E. (2006). *International organizations as law makers*. Oxford, UK: Oxford University Press.
- American Society of International Law. (2006). *International law: 100 ways it shapes our lives*. Retrieved June 30, 2009, from www.asil.org/files/asil_100_ways_05.pdf
- Armstrong, D., Farrell, T., & Lambert, H. (2007). *International law and international relations*. Cambridge, UK: Cambridge University Press.
- Bederman, D. J. (2006). *International law frameworks* (2nd ed.). New York: Foundation Press.
- Birnie, P., Boyle, A., & Redgwell, C. (2009). *Law and the environment* (3rd ed.). Oxford, UK: Oxford University Press.
- Brownlie, I. (2003). *Principles of public international law* (6th ed.). Oxford, UK: Oxford University Press.
- Cassese, A. (2005). *International law* (2nd ed.). Oxford, UK: Oxford University Press.
- Cassese, A. (2008). *International criminal law* (2nd ed.). Oxford, UK: Oxford University Press.
- Claude, R. P., & Weston, B. H. (2006). *Human rights in the world community: Issues and action* (3rd ed.). Philadelphia: University of Pennsylvania Press.
- Cryer, R., Friman, H., Robinson, D., & Wilmshurst, E. (2007). *An introduction to international criminal law*. Cambridge, UK: Cambridge University Press.
- Dinstein, Y. (2004). *The conduct of hostilities under the law of international armed conflict*. Cambridge, UK: Cambridge University Press.
- Donnelly, J. (2003). *Universal human rights* (2nd ed.). Ithaca, NY: Cornell University Press.
- Evans, M. D. (Ed.). (2006). *International law* (2nd ed.). Oxford, UK: Oxford University Press.
- Geneva Conventions. (1949). Available at <http://www.icrc.org/web/eng/siteeng0.nsf/html/genevaconventions>
- Gray, C. (2008). *International law and the use of force* (3rd ed.). Oxford, UK: Oxford University Press.
- Grewe, W. G. (2000). *The epochs of international law*. New York: De Gruyter.
- International Commission on Intervention and State Sovereignty. (2001). *The responsibility to protect: Report of the international commission on intervention and state sovereignty*. Retrieved June 25, 2009, from http://www.iciss.ca/menu_en.asp
- International Committee of the Red Cross. (n.d.). *The mission*. Available at <http://www.icrc.org/HOME.NSF/060a34982cae624ec12566fe00326312/125ffe2d4c7f68acc1256ae300394f6e?OpenDocument>
- International Court of Justice. (n.d.). *Statute of the International Court of Justice*. Available at http://www.icj.cij.org/documents/index.php?p1_4&p2_2&p3_0
- Kennedy, D. (2006). *Of war and law*. Princeton, NJ: Princeton University Press.
- Kolb, R., & Hyde, R. (2008). *An introduction to the international law of armed conflicts*. Portland, OR: Hart.
- Lowe, V. (2007). *International law*. Oxford, UK: Oxford University Press.
- Lowenfeld, A. F. (2008). *International economic law* (2nd ed.). Oxford, UK: Oxford University Press.
- Murphy, S. D. (2006). *Principles of international law*. St. Paul, MN: West.
- Qureshi, A. H., & Ziegler, A. (2007). *International economic law* (2nd ed.). London: Sweet & Maxwell.
- Shaw, M. (2003). *International law* (5th ed.). Cambridge, UK: Cambridge University Press.
- Slaughter, A. (1995). International law in a world of liberal states. *European Journal of International Law*, 6, 503–538.
- Slaughter Burley, A. (1993). International law and international relations theory: A dual agenda. *American Journal of International Law*, 87, 209–239.
- Steiner, H. J., Alston, P., & Goodman, R. (2008). *International human rights in context* (3rd ed.). Oxford, UK: Oxford University Press.
- United Nations. (1945). *Charter of the United Nations*. Available at <http://www.un.org/en/documents/charter>
- United Nations. (1982). *United Nations Convention on the Law of the Sea*. Available at http://www.un.org/Depts/los/convention_agreements/texts/unclos/unclos_e.pdf
- Verheyen, R. (2005). *Climate change damage and international law: Prevention, duties, and state responsibility*. Leiden, the Netherlands: Martinus Nijhoff.

53

INTERNATIONAL ENVIRONMENTAL POLITICS

JERALD MAST

Carthage College

Customers in Beijing and San Francisco restaurants enjoying orange roughly contribute to the collapse of ocean fisheries off New Zealand and Australia. Automobile drivers sitting in Chicago and London traffic add to greenhouse gas emissions that threaten ice sheets in Greenland and Antarctica, which, in turn, may submerge low-lying coastal population centers and displace millions of people from New Orleans to Bangladesh. Beef consumed in Madrid and plywood sold in Tokyo encourage deforestation in Amazonian and Indonesian rainforests, driving species extinction and intensifying climate change. The export of computer waste from France and Canada to China and West Africa results in the release of carcinogenic fumes and heavy metals into rivers and groundwater on which impoverished people depend. The national and international regulations of production and commerce that address or fail to address the consequences of all these activities are shaped by governments, citizens, interest groups, and multinational corporations all over the globe.

In 1971, ecologist Barry Commoner wrote *The Closing Circle*, describing his first and second laws of ecology as follows: Everything is connected to everything else, and everything must go somewhere—nothing goes away. International environmental politics is nothing if not a cautionary tale on these two basic points. The complex, interconnectivity of the modern world in which the human population nears 7 billion persons and global gross domestic

product topped \$65 trillion in 2008 raises important issues of environmental sustainability and social justice that inspire a similarly interconnected web of intellectual and scholarly activity. The subject of international environmental politics involves national, state, and local governments as well as global institutions like the United Nations and World Bank, policies at all levels, markets, technology, interest groups and businesses, and social movements of citizens and consumers. And, of course, it involves the nonhuman world: plants and animals, soil and water, and the intricate processes through which they relate and on which we depend. Finally, the subject of international environmental politics is composed of values, ideas, and philosophies that interpret all of the aforementioned and, on the basis of those interpretations, motivate some actions and discourage others.

The stakes involved in international environmental politics are high. Air and water pollution contribute to the deaths of millions of people of each year; natural resource scarcity endangers the welfare of billions more. The voracious rate of tropical deforestation is a key factor in the current wave of mass extinctions, eroding the planet's biodiversity created by millions and millions of years of evolution. The most dramatic international environmental issue is global warming, which could displace tens of millions of persons, disrupt weather patterns, and alter climate, thus harming agricultural industries across the globe. Understanding the nature of these events and appropriate

responses to them has been the principal objective of scholars of international environmental politics.

Theories of International Environmental Politics

Political scientists are interested in the acquisition and implementation of power on an empirical level—that is, they are interested in how power is acquired and used, but also on a normative level that asks how power should be acquired and used. Some political scientists are more interested in empirical questions, and others are more interested in normative ones, but it is the presence of both forms of inquiry that makes this field of scholarship useful.

How the environment is defined is an important factor distinguishing theories of international environmental politics. Is the environment appropriately understood primarily as a bundle of resources for human use and as a sink for our wastes? Are nonhuman organisms potential commodities to be harvested for our own purposes, disposed of when convenient? Is the environment a resilient, limitless place capable of handling rough treatment? Or should we understand the environment in more reverential terms, as spiritual places, as a fragile network of living systems endowed with intrinsic rights to exist? How the environment is conceptualized is a critical element of the various theories that explain international environmental politics.

Mainstream Theories of International Environmental Politics

The characterizations of international environmental problems typically involve natural resource scarcity, transboundary pollution, and general environmental quality problems. These problems are conventionally described by mainstream theories of political science as common-pool resource problems. Common-pool resource problems are famously described by Garret Hardin (1968) in his essay “The Tragedy of the Commons.” Hardin uses a metaphorical pastoral setting to argue that self-interested individuals seeking to maximize their individual interests will exploit the commons to the point of ecological collapse. Awareness of their collective interests in sustainable use is no help, according to Hardin, for while curtailing your own consumption would benefit the collective good, it would also provide opportunities for competitors to take advantage of your public spiritedness. Thus, altruistic actions on behalf of collective environmental interests impose costs on any individual taking such action, while directly benefiting his or her competitors.

This argument is consistent with public choice theory. Mancur Olson (1965) identifies problems in motivating political action to achieve public goods—goods and services that create nonexcludable benefits. If a policy goal involves a good or service that would create benefits for

everybody (whether you help pay for that good or not), why contribute to the achievement of that policy? The possibility of free riding (enjoying the benefits without paying for them) means that relying on voluntary actions to create public goods like clean air will result in insufficient support. Government regulation is needed to compel support for public goods such as sustainable use of common-pool resources. The alternative, argues Hardin, is to privatize the commons, giving individuals incentives to care for what directly benefits them.

Elinor Ostrom (1990) accepts Hardin’s premise that the incentive structures of common-pool resources encourage individual resource users to overuse the resource and discourage individuals from trying to preserve the resource on their own. But she offers compelling counterarguments to the inevitability of degradation in the absence of coercive regulation or privatization.

Ostrom (1990) finds that ecologically sustainable use of common-pool resources can be voluntarily negotiated by resource users to the extent the following conditions are present among them: (a) a rough consensus about the impending harm from resource exploitation; (b) roughly equal vulnerability to harm; (c) high value placed on sustainable future use of the resource; (d) low information, monitoring, and enforcement costs; (e) a minimal sense of reciprocity and trust exists; (f) relatively low numbers of resource users negotiating exists. The arguments of Hardin and Ostrom reflect different degrees of pessimism or optimism regarding the ability of political actors to negotiate solutions to environmental problems.

Consider these perspectives applied to the global level where the oceans and the atmosphere function as commons and nation-states as individual resource users in an anarchic context. Since the current political arena lacks a global government and is predicated on the concept of national sovereignty, individual nation-states must negotiate solutions among themselves. Studying global politics is the domain of international relations (IR), and the dominant debates within IR on global environmental problems have been consistent with liberal and realist theories of state behavior (Paterson, 2000). In viewing the need for individual states to negotiate solutions to global environmental problems, one sees shades of Hardin-like pessimism in realist theories and shades of Ostrom-like optimism in liberal ones.

Liberal IR scholarship focuses on the emergence and performance of international environmental regimes—defined as a set of negotiated norms, principles, values, regulations, and procedures for decision making to align the interests of states on some international environmental issue. To overcome collective action problems, states collaboratively establish institutional regimes to shape their individual behaviors and serve their mutual interests. Examples of large-scale international collaboration include treaties on acid rain, ozone depletion, whaling, the ban on trafficking in endangered species and ivory, international

trade of toxic substances, biodiversity loss, and climate change (Bryner, 2004). Analyzing international responses to oil spills, acid rain, stratospheric ozone depletion, pollution in the Baltic and North Seas, fisheries mismanagement, farm chemical regulations, and overpopulation, Robert Keohane, Peter Haas, and Marc Levy (1993) find grounds for optimism in the ability of states to negotiate sustainable solutions to global environmental problems.

The degree to which such regimes warrant optimism is a function not only of the facts associated with the particular issues, but also of the analytical perspectives of the observers. Liberals are more committed to diplomacy because, among other reasons, they see it as more possible (please see the Chapter 37, "Realism and Neorealism," and Chapter 38, "Idealism and Liberalism," in this volume). The ends of negotiation—international regimes—are perceived differently as well. Realist theory expects little to come of the regimes; liberal theory has more faith in their efficacy (Paterson, 2000).

A lack of faith in regimes leads many realists to see environmental problems as intractable. This spurs an interest in environmental security. Scholars, particularly realists, are exploring relationships between environmental stresses and the probability of increased rates of social instability, interstate competition, and conflict (Homer-Dixon, 1991). Whether planning for domestic security or for the prospects of international conflict, environmental degradation has increasingly become a legitimate consideration of states.

The debate over nation-states' willingness and capability to collaboratively solve global environmental problems misses important political activities of nonstate actors involved in such issues. Environmental nongovernmental organizations (NGOs) not only target states in order to influence government behavior at legislative and implementation stages of the policy process but also engage in political activity beneath the level of the state, shaping societal norms and attitudes in ways that support their agendas, motivating political action that may or may not require state involvement (Wapner, 1995). The extension of markets, cultural symbols, and social connections across national boundaries has facilitated the emergence of a global civil society where private groups and individuals meet and interact to achieve public ends. Grassroots organizing, public demonstrations, and boycott campaigns are all tools used by transnational environmental NGOs to shift public attitudes and citizen and consumer behavior.

An effective role for transnational environmental NGOs in a global civil society signals a more active role for citizens and consumers than what might be inferred from an otherwise state-centric literature. The theory that democracy has a special relationship with environmental problem solving is based on the history of social movement politics and the passage of environmental legislation. Robert Paehlke (1989) argues environmentalism has evolved to a point where it functions like a distinct progressive ideology

with inherent preferences for decentralized, democratic, participatory decision making. Environmentalism, in other words, has been very good for democracy. Other researchers have argued the inverse, that democracy is inherently conducive to environmentalism (Payne, 1995). That is to say, democracies are far less likely to harm the environment and more willing to take measures to protect it.

The Neo-Malthusian–Anti-Malthusian Debate

British scholar Thomas Malthus (1766–1834) argued that human suffering was inevitable, given that food production only increased arithmetically while population increased exponentially. Although Malthus failed to accurately account for the rate of productivity increases due to industrialized agriculture and the advent of petroleum-based herbicides, pesticides, and fertilizer, his ideas that unchecked population growth would outpace the growth of resource production never really faded and were revitalized in the 1960s and 1970s. During that time, Stanford biologist Paul Ehrlich (1968) wrote *The Population Bomb*, popularizing concerns about environmental and social consequences of exponential population growth.

Four years later, a group of scholars published *The Limits to Growth*, expanding the theme of population by focusing on exponential growth of production and consumption to illustrate the unsustainability of human activity relative to the natural environment (Meadows, Meadows, Randers, & Behrens, 1972). Early ecological economists refined their arguments in the 1970s, led by Herman Daly's (1977) *Steady State Economy*. Daly challenges conventional economic assumptions that reduce the environment to nothing more than discrete subsets (natural resource industries) of the larger economy. Daly argues that it is the economy that is a subset of the larger environment, which provides two critical functions often ignored by conventional economists: The environment is (a) a source of a whole range of necessary materials and energy and (b) a sink into which wastes are put. Both functions are seen as inherently limited by the first two laws of thermodynamics: the conservation of matter and energy and the law of entropy. The implication is that material growth of the economy cannot be sustained in the long run, nor can energy dependence on nonrenewable fuels. Sustainability can be attained, according to ecological economists, by understanding the value of the natural capital base on which societies depend, thus uncovering the real costs of our actions.

But ecological economics has only marginally impacted policy. Certainly economic growth remains a prime objective of governments the world over. During the 1970s, exponentially expanding population and rates of resource consumption and waste production led to dire predictions of ecological catastrophes and political theories about how to avoid them. William Ophuls (1977; Ophuls & Boyan, 1992) and Robert Heilbroner (1974) were prominent

eco-authoritarians expressing profound pessimism about the prospects of achieving sustainability through democratic means.

Ophuls (1977) argues economic growth in liberal democracy is necessary for the political-economic elite to resist demands of the working classes for wealth redistribution. Although numerically few, the elite disproportionately influence policy and won't agree to dramatic wealth redistribution, while the masses, with less power per capita but considerable influence in the aggregate, won't accept freezing current unequal patterns of ownership. Thus, economic growth placates the voting masses with expectations of rising income at little or no cost to the wealthy. What to do about a politically popular but environmentally unsustainable policy? For Ophuls, ecologist-kings are needed to bring the economy to a sustainable steady state, while imposing patterns of income and wealth to maximize political support.

Whether they call for authoritarian responses or not, Dryzek (2005) argues that neo-Malthusians like Ophuls and Meadows (whom he calls survivalists) have the following in common: (a) a deep concern about a finite carrying capacity of the planet for the human species and (b) taking a cue from Hardin, pessimism about human nature's propensity for personal sacrifice for collective ends.

Anti-Malthusians, referred to as Prometheans by Dryzek (2005), reject the basis for neo-Malthusian concerns about environmental sustainability. Humans should be seen as problem-solving resources rather than as sources of environmental degradation. More people means more resources. The most prominent scholars espousing this perspective are Julian Simon and Herman Kahn (1984), who argue that markets generate prices that function as signals of scarcity, creating incentives to (a) conserve or reduce consumption, (b) develop technology to improve efficiency, or (c) find substitutable resources. Thus, trends in human population and economic growth are just part of a larger story of progress, and economic institutions and human ingenuity will solve environmental problems as they arise. Government intervention is generally seen to be unnecessary or counterproductive in that it serves as a drag on economic growth.

These two theoretical perspectives offer starkly different, and in ways dichotomous, interpretations of the world and humanity's relationship to it. However, on a more practical level, these perspectives are poles of a spectrum measuring two variables: capacity of the environment to provide resources and absorb waste and capacity of humans to maximize resources and minimize wastes.

Concern over questions posed by this spectrum fueled interest in theories of sustainable development throughout the 1980s. In 1983, the United Nations created the World Commission on Environment and Development chaired by the prime minister of Norway, Gro Harlem Brundtland, and charged it with addressing the questions of how the international community should respond to environmental degradation and achieve ecological sustainability.

The World Commission on Environment and Development (1987) published a widely read document, *Our Common Future*, which outlined sustainable development as political-economic activity that meets two critical criteria: intragenerational equity and intergenerational equity, defined as development that allows current generations to meet their needs without compromising the ability of future generations to meet theirs. *Our Common Future* affirms a global obligation to address both environmental crises and promote social justice for the world's poor, rejecting Garret Hardin's (1974) so-called lifeboat thesis that aiding the world's poor only exacerbates the unsustainability problem.

At the Earth Summit in Rio de Janeiro in 1992, world leaders expressed support for sustainable development in terms of intra- and intergenerational equity and environmental sustainability as legitimate national and international objectives. Dryzek and Schlosberg (1998) note that widespread support for sustainable development is a function of its ambiguity. They also note that within environmental scholarship, sustainable development has its defenders despite its ambiguity and its critics because of it.

In relation to other ideological frameworks, many within the neo-Malthusian and survivalist position are in line with modern liberal, social welfare-state perspectives, while the anti-Malthusian and Promethean position aligns more comfortably with classically liberal, libertarian thought. Both perspectives, however, share a primary concern for humanity and view the environment in terms of what it provides for societies and economies. A number of increasingly significant theories of the environment and human relationships to it challenge this assumption.

Radical Theories of International Environmental Politics

Robyn Eckersley (1992) argues that certain elements of environmentalism have challenged Western political thought by attacking the premise of anthropocentrism (the notion that the value of the natural world resides in its utility to human beings) and offering in its place ecocentrism, the belief that all life, and the ecological systems on which life depends, possess intrinsic value. Environmentalism associated with mainstream politics and with neo-Malthusian arguments equates the environment with natural resources, focusing on the need to manage these resources in equitable and sustainable fashions. To the extent some environmentalists think in these terms, Eckersley contends, they merely modify various forms of conventional political theory. Where environmentalism is truly transformative is where it takes seriously the idea of inherent value in places and nonhuman life and envisions how we might order societies in such ways that acknowledge and respect that.

Deep ecology is one such theory. Largely based on the writings of Norwegian philosopher Arne Naess (1973), deep ecology challenges the necessity and appropriateness of conceiving the self in individual terms. Instead, deep

ecologists call for a recognition that, individually, we are all enmeshed in and dependent on social and ecological communities and that our identities—our selves—ought to reflect those connections and dependencies. The personal and psychological transformations that come with this redefinition of self will facilitate a more sensitive, peaceful, and nonhierarchical transformation of our relations to each other and the natural world.

This notion of hierarchy as the dynamic behind environmental degradation is shared by ecofeminists. For deep ecologists, the hierarchy is anthropocentric: humans dominating the natural world. For ecofeminists, the hierarchy is also patriarchal: men dominating women (Warren, 1996). These forms of domination are not mutually exclusive, and ecofeminists argue that they reinforce each another in a multitude of ways. Warren identifies these forms of association between women and nature as historical, conceptual, empirical and experiential, epistemological, symbolic, ethical, theoretical, and political. Women and nature, according to various versions of ecofeminist thought, are consciously as well as unconsciously associated with each other in ways that rationalize their subjugation to men and male-dominated societies. Effectively responding to environmental crises requires an understanding of the ways men perceive and mistreat both women and nature and altering these sexist belief systems and the patterns of behavior they attempt to justify. Like deep ecology, the basic end for the many versions of ecofeminist thought is nonhierarchical societies based on respectful, reverential senses of reciprocity and mutuality between human and nonhuman life.

There are many other radical environmental thinkers, including bioregionalists, social ecologists, eco-socialists, and eco-anarchists. They are united by a desire to transform social systems to create “new patterns of production, reproduction and consciousness that will improve the quality of human life and the natural environment” (Merchant, 1992, p. 9). They possess preferences for grassroots, decentralized approaches to political action. With a local focus and a preference for changing social attitudes within civil society rather than engaging largely unredeemable states, many radical environmental thinkers end up ceding key portions of the international arena to more reform-minded theorists. But because they also refuse to accept as legitimate forms of value and thought that they see as inherently oppressive, political-economic institutional processes based on those forms of thought tend to exclude them as well.

A Theoretical Synthesis? Ecological Modernization

What makes mainstream theories of international environmental politics mainstream is the extent to which they accept the suitability of preexisting political-economic institutions and prevailing cultural and philosophical

viewpoints to address global environmental problems. Where one comes down on this question of suitability is really a function of two factors: one's perception of the sustainability question addressed by the neo-Malthusian–Anti-Malthusian debate and the degree of faith one has in both the political-economic institutions and the dominant cultural values themselves. One may seriously believe in the inability of ecosystems to bear much more stress from expanding levels of human population and consumption and still believe that the bases of preexisting political-economic institutions are (a) worth saving in their own right through environmentally minded reformation or (b) inevitable and therefore must be so reformed since no acceptable alternative awaits.

The theory of ecological modernization emerged in Germany, the Netherlands, and Britain during the 1980s and takes seriously the belief that the world's ecosystems are limited and that carrying capacities are real and constitute boundaries that societies cross at their peril. But ecological modernization theorists also maintain that industrial capitalist societies can reform their central institutions and practices so as to achieve conditions that are environmentally just and sustainable. By enlisting the aid of businesses to develop new technologies and renewable energies, ecological modernization theory expects capitalist societies to improve efficiency and lower waste and emissions, simultaneously increasing wealth while reducing harm to the environment. This will involve active governments committed to environmental protection in all that they do, primarily by working in partnership with business to assist industry and commerce to incorporate ecological ideas. A key assumption of ecological modernization theory is that collaboration and negotiation are more effective at achieving a sense of common purpose, and thus more capable of constructing coalitions willing to support and strong enough to enact the kinds of policies to restructure society along ecologically sustainable lines (Hajer, 1995).

Applications and Empirical Evidence

Measuring the success or failure of international environmental policies is complicated for empirical and theoretical reasons. As an empirical matter, environmental politics cover wide-ranging and complex subjects, and understanding current conditions of large-scale ecosystems requires longitudinal data that are often incomplete or nonexistent. When such data exist, their interpretation invariably involves scientific uncertainty, which is compounded when modeling future consequences. Latent effects and the potential for unforeseen negative and positive feedback loops make assessment of current conditions difficult and accurate predictions of future outcomes of policies even more so.

Assessing the desirability of policy outcomes involves formal or informal comparison of benefits and costs, which are often hard to quantify. How should we measure

the value of lives saved, illnesses reduced, and beautiful and biologically diverse places preserved? Techniques exist to measure such things but are methodologically limited and conceptually controversial. Problem displacement is another complicating factor; resolving one particular problem often involves shifting harm to other areas. For example, the Oslo and Paris Conventions created an international regime to reduce pollution in the North Sea. To comply with restrictions on ocean dumping, Britain increased land-based dumping and incineration, reducing ocean pollution but raising levels of land and air pollution (Skjaereth, 1998). Given the complexity of the contexts in which international environmental problems occur, it is also often difficult to ascribe causality to a given policy when multiple ecosystem and jurisdictional boundaries are involved.

Theoretically, there is no consensus on the nature of environmental problems that would allow us to objectively define success. The criteria by which we define it involve an assortment of normative assumptions. If one begins with an ecocentric assumption that nonhuman life and ecosystems deserve respect in their own right independent of their utility to humans, it is very unlikely that one will perceive much progress in international environmental issues.

These difficulties aside, some consideration of global environmental issues can shed light on the efficacy of international environmental policies. Research suggests there is a good deal of variation in the rate of success or failure for international environmental treaties (see Chapter 51, "International Law," for an overview of the associated challenges such efforts face). Some regimes are successful, others much less so. The Antarctic Treaty System took effect in 1961, successfully preventing international conflict over the continent of Antarctica and facilitating scientific research there (Young, 1997). The Montreal Protocol on Substances that Deplete the Ozone Layer was passed in 1989 to reduce emissions of industrial and commercial chemicals that break down stratospheric ozone, which blocks harmful ultraviolet solar radiation. Although challenges in developing nations and in black markets persist, the Montreal Protocol is successfully addressing this global environmental problem, having achieved a 95% reduction in use of ozone-depleting chemicals and resulting in tangible ozone recovery (Parson, 2003).

Sulfur dioxide and nitrogen oxide emissions create acid rain, adversely affecting aquatic and forest ecosystems, crops, and buildings. In the 1970s, there were no regulations specifically addressing acid rain, and the use of higher smoke stacks to improve ambient air quality may have increased the problem (Munton, Sooros, Mikitina, & Levy, 1999). In 1979, led by Sweden, Norway, and Finland, 33 states agreed to the Convention on Long-Range Transboundary Air Pollution (LRTAP). Lacking specific rules or prescriptions because of opposition

from the United States, Britain, and Canada, the treaty nevertheless contributed to behavior changes in its signatories (Munton et al.). The full extent to which the acid rain problem has been solved is debatable, but sulfur dioxide emissions in North America and Europe have declined markedly, as have, to a lesser extent, nitrogen oxides.

Many international regimes to protect ocean fisheries have been failures. But one of the more successful regimes in inducing behavior change in its membership is the International Whaling Commission (IWC) created by the International Convention for the Regulation of Whaling in 1946 (Andresen, 1998). Originally an instrument to promote whaling, the IWC passed policies that encouraged intensive whaling practices in the 1940s and 1950s, severely depleting many whale populations. In the 1970s, however, NGOs like Greenpeace made the whale an icon of environmentalism, successfully increasing public opposition to whaling. Antiwhaling states swelled the ranks of the IWC and are transforming it into an instrument for whale preservation, evidenced by the moratorium on commercial whaling passed in 1982. Although Japan continues to practice so-called scientific whaling, overall the commercial catch of whales has declined dramatically (Andresen).

Progress addressing the most prominent global environmental problem—global warming—has been much harder to achieve. Greenhouse gas (GHG) emissions like carbon dioxide and methane trap solar radiation within Earth's atmosphere as it radiates from the surface. The increased composition of greenhouse gases is correlated with rising global temperatures and shifting climate patterns. The composition of GHG in the atmosphere has fluctuated naturally throughout Earth's history, but since the beginning of the Industrial Revolution, anthropogenic contributions from the combustion of fossil fuels have contributed significantly.

In 1988, the United Nations Environment Programme and the World Meteorological Organization established the Intergovernmental Panel on Climate Change (IPCC) to assess the scientific literature on climate change in order to provide guidance to political leaders and the public about causes and consequences of and possible responses to global climate change. As an agency for gathering and publicizing information, the IPCC has been successful, issuing major assessment reports in 1990, 1995, 2001, and 2007 and winning the Nobel Peace Prize in 2007. Moreover, the first IPCC report served as the basis for the United Nations Framework Convention on Climate Change (UNFCCC) written at the Earth Summit at Rio de Janeiro in 1992.

The goal of the UNFCCC was to stabilize anthropogenic greenhouse gas emissions at safe levels. It created a national GHG emissions inventory to track the amount of emissions and a nonbinding commitment by industrialized nations to stabilize GHG emissions at 1990 levels by 2000.

The UNFCCC divided signatory states into three categories: developing states that would not be expected to meet the 1990 emission targets, developed states that would, and a subset of developed states that would financially assist developing states in implementing emission-reducing technologies.

In 1997, the signatory states adopted the Kyoto Protocol, amending emission reduction targets to 5% below 1990 levels by 2008 to 2012 for industrialized countries collectively (national targets vary) and making them legally binding. Among industrialized nations, only the United States (until 2006 the world's largest emitter of carbon dioxide) has refused to ratify the Kyoto Protocol. Despite widespread participation, however, anthropogenic carbon dioxide emissions grew 4 times faster after 2000 than in the 1990s. Although growth of emissions from industrialized nations slowed, they did grow, and emissions in developing nations, particularly China and India, grew rapidly. These developments support critical assessments of the international environmental regimes on climate change (Victor, 2002).

Research assessing potential consequences of global climate change notes the likely displacement of 200 million persons due to rising sea levels, rising death rates from malnutrition and heat stress, increasing flooding during wet seasons, drought during dry seasons, and the extinction of 15% to 40% of the planet's species if warming trends go unabated (Stern, 2007). Costs of global warming have largely been viewed in terms of future impacts, but increasingly researchers are identifying current harms. One recent report estimates that already 300,000 people per year die of causes related to climate change, and \$125 billion in average annual economic costs occur as well (Global Humanitarian Forum, 2009). Although these estimates are approximations reached with controversial methodologies, it is unequivocal that many people in the world today are harmed and die as a result of environmental problems, and many others lead tenuous lives made vulnerable by degraded environmental conditions.

The UN's Human Development Programme estimates more than 1 billion people lack regular access to safe drinking water; one result is that 1.8 million children die each year from treatable and avoidable diseases caused by unsanitary water (Watkins, 2006). Comparable numbers of people suffer from chronic hunger due to food insecurity. Outdoor air pollution causes more than 2 million premature deaths annually. Elevated cancer rates are associated with cities, hazardous waste sites, areas downwind from certain industrial activities, and farming states and are linked to known carcinogenic substances in solvents, petro-chemicals, dioxins and other industrial by-products, and pesticides. Although global in nature, costs created by pollution are, like natural-resource vulnerability, not experienced equally. The poor within and across nations are at greatest risk, underscoring how little sustainable development has actually been achieved and how much the

aspirations of environmental quality and equity espoused by *Our Common Future* have gone unmet (Sneddon, Howarth, & Nordgaard, 2006).

Policy Implications and Future Directions

Given the mixed record of international political efforts to resolve global environmental problems, what lessons can be drawn about how to craft effective environmental regimes? If the larger picture of overall global environmental conditions is as troubling as much of the evidence suggests, what does that imply for larger political-economic systems? This section addresses these questions in brief.

Research by Young (1997) and Victor, Raustiala, and Skolnikoff (1998) identify factors that account for the range in effectiveness of international environmental regimes. Such regimes are more likely to effectively change state behavior in positive ways to the extent that Pareto improvement (win-win, or at least win-not lose) situations have been identified and agreed on (Young). Regimes in situations with zero-sum outcomes are far more likely to be resisted or subverted by likely losers. One potentially useful response is the use of nonbinding targets, which may put states facing considerable compliance costs more at ease. One example is the LRTAP convention to reduce acid-rain-causing pollutants discussed earlier (Munton et al., 1999). A potential danger at the stage of defining targets is states purposely setting conservative targets that are easily met, allowing them to achieve legal compliance without significantly addressing the problem. This danger is less likely when there is transparency in the policy process, the problem to be addressed is salient, and high levels of public involvement exist (Victor et al., 1998).

Flexibility in how regime goals are met creates significant advantages in that it allows states with unique circumstances to take those circumstances into account in crafting policy approaches, yet when flexibility leads to a diversity of policy responses, the evaluation of the regime as a whole is made more complicated. Evaluation is important, for it helps identify situations where states willfully neglect their obligations, as well as situations where states have poorly planned or miscalculated (Victor et al., 1998). The combination of flexibility in rule making and good evaluation data allows states to practice adaptive management, which is taking corrective actions as new conditions emerge and understanding of issues improves (Young, 1997).

International environmental regimes are more likely to be effective when there are active domestic constituencies within the individual states who have a stake in the success of the regime (Young, 1997). When these constituencies involve both the industry targeted by the regime and environmental nongovernmental organizations, opportunities

to accumulate data on implementation and negotiate mutually acceptable policy amendments are improved (Victor et al., 1998). States with domestic constituents with interests in regime effectiveness are more likely to be pressured into actions that will bring them into compliance.

These factors for international environmental regime effectiveness are consistent with other observations on the evolution of environmental policy more generally. Theorists and activists have begun to advocate shifts away from conventional, command-and-control regulatory responses to environmental problems where government sets a particular standard of behavior and compels regulated parties to meet it. Instead, more pragmatic and collaborative approaches to problems in which stakeholders are allowed to negotiate over a variety of solutions ranging from public participation to the inventive use of property rights are seen as generally more effective and efficient in achieving results (Durant, Fiorino, & O'Leary, 2004). The aforementioned principles may be useful in crafting better, more effective international environmental regimes in the future.

A range of such policy tools to address climate change have been developed, the most prominent being cap-and-trade systems, carbon taxes, and carbon offsets. All are potentially useful and will play roles in national and international responses to climate change. A real sticking point in negotiating new protocols will be how to treat different states in terms of their obligations to reduce emissions. Emissions from China and India have grown faster than those of the rest of the world. Yet they are poorer and less developed, and their leaders argue the developing world deserves to advance its economic conditions, as did the developed, industrialized states. Moreover, because greenhouse gases are persistent in the atmosphere, much of the problem has already been created by industrialized developed states. Many critics of industrialized states ask why developing states should bear equal responsibility for fixing a problem that they played only a minor role in creating.

The theory of ecological modernization seems likely to play a prominent role in decades to come. It has, however, been mainly a European phenomenon, both in theory and in practice. What can it offer developing nations? Arthur Mol (2003) argues that although economic globalization has led to widespread instances of environmental degradation, it also has created situations where environmental outcomes have improved. Regional integration can raise environmental standards in laggard states, where a multitude of nonstate actors, NGOs, popular social movements, and multinational corporations can foster environment-oriented discourse within civil society at local, national, and global scales. Identifying instances where globalization has led to environmental improvements and reduced poverty in the developing world is of critical importance given its overall disheartening track record.

Conclusion

The scale and scope of the human presence on Earth has reached a point where the aggregate consequences of both appear to be affecting the ability of the planet to continue to sustain us and our actions for much longer. Our presence, in terms of population and economic activity, has heightened the interconnectivity of the modern world. Our actions no longer, if they ever did, exist in isolation. Issues of ecological sustainability and equity—intragenerationally, intergenerationally, and across species—all involve compelling questions and challenges.

Can competitive and self-interested states sustainably manage common-pool resources? Are neo-Malthusian fears of a surpassed planetary carrying capacity valid? Should anti-Malthusian faith in human ingenuity, markets, and technology be enough to allay those fears? If not, can liberal capitalism be sufficiently greened? A concern for the welfare of humanity motivates religion and philosophy and serves as the guiding passion of much environmentalism. But a passionate concern for humanity can mask an exclusive concern for humanity, and in viewing the world and its life forms around us as mere resources, we may promote an abuse of the natural world that not only does an injustice to the rights and interests of nonhuman life but also may undermine our own welfare in the process. The academic field of international environmental politics tries to understand these questions and issues, identify possible answers to them, and shape effective solutions on the basis thereof.

References and Further Readings

- Andresen, S. (1998). The making and implementation of whaling policies: Does participation make a difference? In D. G. Victor, K. Raustiala, & E. B. Skolnikoff (Eds.), *The implementation and effectiveness of international environmental commitments: Theory and practice* (pp. 431–474). Cambridge: MIT Press.
- Barry, J., & Eckersley, R. (Eds.). (2005). *The state and global ecological crisis*. Cambridge: MIT Press.
- Bryner, G. C. (2004). Global interdependence. In R. F. Durant, D. J. Fiorino, & R. O'Leary (Eds.), *Environmental governance reconsidered: Challenges, choices, and opportunities* (pp. 69–104). Cambridge: MIT Press.
- Daly, H. E. (Ed.). (1972). *Toward a steady state economy*. New York: W. H. Freeman.
- Daly, H. E., & Cobb, J. B. (1994). *For the common good: Redirecting the economy toward community, the environment, and a sustainable future*. Boston: Beacon Press.
- Dryzek, J. (2005). *Politics of the Earth*. New York: Oxford University Press.
- Dryzek, J., & Schlosberg, D. (Eds.). (1998). *Debating the Earth: The environmental politics reader*. New York: Oxford University Press.
- Durant, R. F., Fiorino, D. J., & O'Leary, R. (Eds.). (2004). *Environmental governance reconsidered: Challenges, choices, and opportunities*. Cambridge: MIT Press.

- Eckersley, R. (1992). *Environmentalism and political theory: Towards an ecocentric approach*. Albany: SUNY Press.
- Ehrlich, P. R. (1968). *Population bomb*. New York: Ballantine.
- Global Humanitarian Forum. (2009). *The anatomy of a silent crisis*. Retrieved May 31, 2009, from <http://assets.ghf.ge.org/downloads/humanimpactreport.pdf>
- Hajer, M. A. (1995). *The politics of environmental discourse: Ecological modernization and the policy process*. Oxford, UK: Clarendon Press.
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162, 1243–1248.
- Hardin, G. (1974). Living on a lifeboat. *Bioscience*, 24, 561–568.
- Heilbroner, R. (1974). *Inquiry into the human prospect*. New York: W. W. Norton.
- Homer Dixon, T. F. (1991). On the threshold: Environmental changes as causes of acute conflict. *International Security*, 16(2), 76–116.
- Keohane, R. O., Haas, P. M., & Levy, M. A. (1993). *Institutions for the Earth: Sources of effective environmental protection*. Cambridge: MIT Press.
- Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. H. (1972). *The limits to growth*. New York: Universe Books.
- Merchant, C. (1992). *Radical ecology: The search for a livable world*. New York: Routledge.
- Mol, A. J. P. (2003). *Globalization and environmental reform: Ecological modernization of the global economy*. Cambridge: MIT Press.
- Munton, D., Sooros, M., Nikitina, E., & Levy, M. (1999). Acid rain in Europe and North America. In O. Young (Ed.), *The effectiveness of international environmental regimes* (pp. 155–247). Cambridge: MIT Press.
- Naess, A. (1973). The shallow and the deep: Long range ecology movement: A summary. *Inquiry*, 16, 95–100.
- Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Ophuls, W. (1977). *Ecology and the politics of scarcity: Prologue to a political theory of the steady state*. San Francisco: W. H. Freeman.
- Ophuls, W., & Boyan, A. S. (1992). *Ecology and the politics of scarcity revisited: The unraveling of the American dream*. New York: W. H. Freeman.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. New York: Cambridge University Press.
- Paehlke, R. C. (1989). *Environmentalism and the future of progressive politics*. New Haven, CT: Yale University Press.
- Parson, E. A. (2003). *Protecting the ozone layer: Science and strategy*. New York: Oxford University Press.
- Paterson, M. (2000). *Understanding global environmental politics: Domination, accumulation and resistance*. New York: St. Martin's.
- Payne, R. A. (1995). Freedom and the environment. *Journal of Democracy*, 6(3), 41–55.
- Simon, J. L., & Kahn, H. (1984). *The resourceful Earth*. New York: Blackwell.
- Skjaerseth, J. B. (1998). The making and implementation of North Sea commitments: The politics of environmental participation. In D. G. Victor, K. Raustiala, & E. B. Skolnikoff (Eds.), *The implementation and effectiveness of international environmental commitments* (pp. 327–380). Cambridge: MIT Press.
- Sneddon, C., Howarth, R. B., & Nordgaard, R. B. (2006). Sustainable development in a post Brundtland world. *Ecological Economics*, 57, 253–268.
- Stern, N. (2007). *The economics of climate change: The Stern review*. Cambridge, UK: Cambridge University Press.
- Victor, D. G. (2002). *Collapse of the Kyoto Protocol and the struggle to slow global warming*. Princeton, NJ: Princeton University Press.
- Victor, D. G., Raustiala, K., & Skolnikoff, E. B. (Eds.). (1998). *The implementation and effectiveness of international environmental commitments*. Cambridge: MIT Press.
- Wapner, P. (1995). Politics beyond the state: Environmental activism and world civic politics. *World Politics*, 47, 311–340.
- Warren, K. J. (1996). *Ecological feminist philosophies*. Bloomington: University of Indiana Press.
- Watkins, K. (2006). *Human development report 2006: Beyond scarcity: Power, poverty and the global water crisis*. Retrieved May 31, 2009, from <http://hdr.undp.org/en/reports/global/hdr2006/>
- World Commission on Environment and Development. (1987). *Our common future*. New York: Oxford University Press.
- Young, O. (Ed.). (1997). *Global governance: Drawing insights from the environmental experience*. Cambridge: MIT Press.

PART IV

POLITICAL SCIENCE METHODOLOGY

54

EVOLUTION OF SCIENCE IN POLITICAL SCIENCE

JAMES K. L. WONG

London School of Economics and Political Science

The scientific study of politics bears a rather short history. It was not until the 1950s that political science reached its attic as a distinct academic discipline. The less-than-a-century time frame, however, has seen significant developments in terms of theoretical and methodological divides. From positivism and interpretivism before the 1980s to a synergy of both thereafter, each of these prominent paradigms not only advocates different approaches to political analysis but also shares varying assumptions about the science of social inquiry. This chapter offers a general overview of the evolution of science and scientific methods. The central questions addressed include the following: (a) What is science and how can the study of politics be scientific? and (b) How did the contemporary debates in the philosophy of (social) science shape the methodological development in political science?

Conceptualizing Political Science: Truth, Knowledge, and Scientific Method

In its simplest sense, political science is a subbranch of social science, which means the scientific study of the political in human society, such as political actions, systems or institutions, and outcomes. This is often conducted through constructing concepts, models, and theories in order to describe and explain what the properties of the political and the underlying causal mechanisms are. Such

inquiry of politics is in sharp distinction from the philosophical tradition of the discipline, which instead seeks to evaluate the political on the basis of normative values and principles. Although the latter tradition, which originated in ancient times, shares a much longer history, the former has expeditiously occupied a dominant standing since the mid-20th century. How did the distinctive tradition evolve from there to here over the last several decades? One must begin with what political science is all about.

Literally, political science is composed of two essential components. The concept of the political is essentially contestable, but roughly it refers to any matters that involve or affect two or more individuals in which disagreements may arise. Under this definition, the political takes place in any collective human entity as a process of conflict resolution, which is composed of interactions among the agents within (action), arrangements of the entity per se (system), and decisions that form the resolution (outcome). An example is the application of an electoral college (system) in choosing a candidate to be the leader of a country (outcome) based on preferences of the voters within. These are the basic dimensions that the study of the political seeks to examine. Then what does it mean to approach these dimensions with a science? Likewise, the concept is in no way self-explanatory, and it is open to various interpretations and controversies (e.g., see Popper's distinction between science and pseudoscience in a following section), yet it is essential to offer a minimal working definition. According to the *Oxford English*

Dictionary, the term *science* can be defined broadly as “knowledge or cognizance of something specified or implied” or narrowly as “a branch of study which is concerned . . . with a connected body of demonstrated truths . . . and which includes trustworthy methods for the discovery of new truth within its own domain.” In this way, science is both the content and the method—it is a representation of truth in the form of knowledge plus the way that such knowledge is acquired. For example, chemistry aims at understanding (knowledge) the properties, structure, and composition of matter in nature (truth) through experimentation, observation, and theory construction (method). Political science, similarly, is the use of the scientific method in obtaining knowledge of truth in politics.

An immediate question would be in what ways scientific inquiry in politics is comparable to that in natural sciences such as chemistry and physics. This concerns whether there is any truth in human society as there is in the natural world, and if so whether it is possible and what method can be used to acquire the knowledge of it scientifically. To start, what are truth, knowledge, and scientific method? Take, as an example, the celebrated Duverger’s law in political science, a principle that holds that proportional representation and a majority vote on two ballots tend toward a multiparty system (Duverger, 1972). One would, without hesitation, regard Duverger’s law as providing a kind of knowledge of the political in terms of the relationship between electoral systems and outcomes. Such knowledge is represented in the form of a principle, or in more formal language a *proposition*, where *proportional representation*, *majority vote on two ballots*, and *multiparty system* are concepts, or components of the proposition. These concepts per se, however, are seldom considered knowledge unless they are being used in a proposition. For example, the term *proportional representation* itself carries no substantial meaning unless it is placed in a proposition, such as the following: Proportional representation means that the percentage of votes obtained by a group of candidates approximates that of the seats allocated to the corresponding group in an election. The above conceptualization of proportional representation, nevertheless, may still not be knowledge, but it is a necessary condition for that concept to be situated in a proposition to qualify as knowledge. But what determines whether a proposition is knowledge? This concerns the nature of knowledge.

When a person, *i*, says he or she knows the meaning of proportional representation as in the above proposition, for instance, he or she must at least believe that the proposition is true. However, a true belief does not exhaust the criteria of knowledge. Suppose *i* (mistakenly) believes that proportional representation and a majority vote on one ballot are two identical types of election systems; *i* can still assert that the proposition is true, although that proposition cannot be considered knowledge at all. This is because *i* fails to justify why the proposition is true, and at most he or she is said to believe, rather than know, it to be true. In other words, knowledge, as widely agreed in epistemology, is justified

true belief. One interpretation of justification is the use of reasons in substantiating a belief to be true. For example, one can analyze the meaning of proportional representation by arguments to demonstrate why the proposition in the previous paragraph, instead of the one believed by *i*, is true. Meanwhile, one can also test and prove Duverger’s law by collecting empirical evidence from the real world (e.g., Riker, 1986). In both cases, the true propositions as confirmed are knowledge, and hence science, whereas the reasoning applied in arriving at such confirmation is also science, or more precisely, scientific method.

However, why does scientific method lead to (scientific) knowledge? Even though one may take a scientific approach (e.g., experiments and evidence collecting), the proposition under investigation cannot be concluded true right away without establishing a relation between the truth and the results drawn from the scientific method. There are two rival accounts of truth—namely, the correspondence theory and the coherence theory. In a nutshell, the correspondence theory of truth posits that a proposition is true if and only if it corresponds to a relevant (set of) fact(s) (e.g., David, 2008), whereas the coherence theory of truth alleges that a proposition is true if and only if it coheres with a relevant reference set of propositions (Young, 2008). For instance, according to the former theory, Duverger’s law can be true if and only if it represents the corresponding fact of the casual correlations between the proportional representation or majority vote on two ballots and a multiparty system. On the other hand, according to the latter theory, Duverger’s law can still be true even if it does not represent such a fact but instead is consistent with, or at least implied by, a certain specified set of propositions. The specified reference set is assumed to be true, which can be defined by the set of propositions currently believed by actual people (Young, 1995) or which will be believed when people have reached a certain level of inquiry (Putnam, 1981). For example, suppose in a particular space and time, people insist that there is a negative correlation between proportional representation and a multiparty system; Duverger’s law, according to the coherence theory, will be judged false irrespective of whether it may be otherwise in concordance with other criteria of truth. The divergence between the two views on truth is not at all subtle, nor does it belong to a purely epistemological issue in philosophy. It is, instead, fundamental to the methodological debate in both contemporary natural and social and political sciences.

What is the origin of the methodological disagreement? Why is there such a disagreement? Why do the two contending positions of truth affect the evolution of science (or more precisely scientific method) in social and political science? This traces to the very basic assumptions of the two theories of truth, to wit, (scientific) realism and antirealism. According to realism, the entities depicted by any true (social) scientific propositions do, by definition, objectively exist, albeit they may not be directly observable, and such propositions are the best (approximate) descriptions of these entities. For example, for realists, the

positive causal correlations between proportional representation and a multiparty system in Duverger's law is a piece of social fact, given that the law (or proposition) itself is scientifically proven true. The realist position is generally grounded on the no miracles argument where the success of a (social) scientific theory, say the explanatory and predictive power of Duverger's law as applied in many real cases of election, would be simply a mystery if there were no objectively true entities and regularities in the world (Putnam, 1975). On the other hand, antirealists deny the objective existence of such entities as (social) facts, and one eminent argument is that these entities only exist to the extent to which they are shaped by human knowledge and minds (e.g., Kuhn, 1962). Antirealists would judge Duverger's law to be true, in accordance with, for instance, the coherence theory of truth, only when it is coherent with what is believed to be true through human understanding. Given the divergent positions of the existence of (social) facts of realists and antirealists, the next question is whether they also differ in terms of the scientific methods used in arriving at those true scientific propositions. The following section illustrates how the realism-antirealism debate molds the evolution of science in terms of scientific methods in social and political science.

Evolution of Political Science: Debating the Scientific Methodologies

Beginning: Positivism and Inductive Political Science

Positivism presupposes a realist standpoint that posits that (social) facts objectively exist, and are external from humans and independent of human knowledge and understanding. For positivists, a proposition is true, and hence constitutes knowledge, if and only if it is in a correspondence relation with the objective fact(s). Since the facts are out there in the world, it is necessary that the reasoning that reveals such correspondence relation (if any) must go beyond pure formality and extend to reality. Among the Vienna Circle in the 1920s, a group of logical positivists attributed the distinctiveness of scientific knowledge to its possible derivation from the facts of experience in the real world. They distinguished between two kinds of propositions—namely, analytic and synthetic. The former are deductively provable through logical reasoning alone, while the latter are empirically verifiable through logical reasoning by reference to empirical facts obtained through the use of the senses and capture what the positivists regard as knowledge. But how is it possible to reason with empirical facts? This is done by inductive logic. In principle, the positivists see scientific knowledge as a form of universal generalization from empirically observable facts. In the natural sciences, for instance, Newton's first law of motion, which states that any physical object remains static or in uniform motion unless being acted on by an external force (or inertia), represents

a body of scientific knowledge that generalizes the motional properties of physical objects. This knowledge can be obtained by observations, such as through experiments, on a large number of instances in which all entities being observed possess a certain property (inertia) under some conditions (the absence of external force), and an inference can be drawn that all entities (physical objects), even those which have not been observed, exhibit inertia when subject to no external force. Such an inference is a universal generalization arrived at by inductive logic in this form: The property of many (or even some) implies the property of all. It is not difficult to spot that such reasoning is not necessarily valid at all times—it can be true that many entities possess a certain property while there is an unobserved entity that behaves the other way around and thus renders the conclusion false. This concerns the idea of falsification and will be addressed in further detail in the next section.

Positivists in the social sciences (including political science) believe that scientific knowledge is also obtainable in the form of universal generalizations through inductive inference similar to that used in the natural sciences. Assuming the existence of objective social facts, through observations and inductive reasoning, these bodies of social facts can be identified and generalized in terms of both taxonomy and relations. The following two statements are propositions of taxonomy identification, and Duverger's law generalizes the relations between both taxonomies mentioned: (1) A significant number of countries (n) operate proportional representation in their parliamentary elections and (2) the parliaments of almost all of the above countries (e.g., $n - 1$) are composed of multiple parties. It is imperative that, to positivists, the social facts be objective and observable in order to be represented as knowledge, and hence, normative questions, such as whether proportional representation should be adopted, are excluded from what can be counted as knowledge that cannot be detected through observations or determined without subjectivity.

Behavioralism has much appeal to positivism and inductive reasoning. In (scientific) methodology, it took a leading position in the 1950s and 1960s when the discipline of political science started to become popular.¹ As implied by its name, behavioralism develops scientific knowledge of the political by focusing on phenomena involving either individual human behaviors (the microlevel) or their social aggregate (the macrolevel) in the political process.² First, it presupposes the realist standpoint that there exist bodies of objective and quantifiable social facts in the political arena that represent the truth—that is, political phenomena (and their mutual relations) do exist and exhibit certain regularity over time. Second, similar to positivism, the truth is observable so that empirical evidence can be collected to identify (the relations of) these political phenomena. These (causal) relations of phenomena can be further generalized, usually through vigorous statistical analysis, comparative case studies to formulate scientific propositions or theories through inductive reasoning, or both. For behaviorists, these

propositions and theories correspond to the truth that is there in the observable political realm, and they are constructed objectively, based purely on empirical evidence without any a priori assumptions or subjective human interpretation. Behavioralists believe that their approach to political science is genuinely scientific, and the propositions and theories so formulated offer a systematic explanation of past and present political phenomena as well as a well-grounded prediction of those in the future. Examples of seminal works include Truman (1951), Lipset (1959), Dahl (1961), Duverger (1964), Parsons (1966), and Lijphart (1971).

Is the picture described a complete one? Does positivism provide a satisfactory account of truth? Does induction constitute the best scientific method in revealing the truth as knowledge? Can behavioralism be the only scientific approach in political science? The historical development of (political) science shows that answers to these questions are not necessarily positive, as a result of the received critiques of positivism, inductivism, and behavioralism by their counterparts.

Adjustment: Falsificationism and Deductive Political Science

Consider again that inductive reasoning does not guarantee true conclusions (generalizations) from true premises (observations). Pinpointed by the 18th-century Scottish philosopher David Hume (1711–1776), the problem of induction lies in the fact that it fails to construct any generalization of empirical observations that can be proven to be true. This matters particularly when the generalization is supposed to make more or less accurate predictions about the future. The sunrise example can illustrate this point: Even if the sun rises every single day as observed in the past, this does not mean that the same will happen every single day in the future. The common sense that the sun rises every day assumes what Hume called the “uniformity of nature.” This assumption contends that the objective truth exhibits absolute regularity, and the possibility of variation is ruled out. Therefore, even if the sunrise tomorrow is not yet observed (and unobservable here and now), inference can still be drawn to predict that the sun will rise tomorrow based on empirical evidence that the sun did rise every day in history. However, there is no reason that the uniformity of nature assumption must be true in the first place, since subsequent discontinuation of regularity (i.e., change) is at least theoretically possible as long as it is not (yet) proven to be otherwise—that is, there is no way that a proposition or theory is verifiable by inductive reasoning.

Given the possible loopholes of inference by inductive logic, the determination of whether a proposition corresponds to the truth (assuming the existence of such) and hence the representation of a body of scientific knowledge cannot rely on verification as the criterion. This is because the correspondence relation between the proposition and the truth may cease to hold should any change occur about the truth in the future. A proposition that captures the relevant

social facts is even more vulnerable to such query because social and political changes are by no means rare. Therefore, an alternative standard is necessary for identifying propositions that constitute (scientific) knowledge or another conception of the correspondence relation between the truth and the proposition. Consider Duverger’s law again. Suppose the correlation between proportional representation and a multiparty system exists as social fact and all empirical evidence confirms the law at present, and there would be a political regime R after n years where its adoption of proportional representation leads to a two-party, instead of a multiparty, system. Such a single instance (R_n) would be sufficient for arguing that Duverger’s law does not always hold. To use a more technical term, R_n would falsify Duverger’s law, provided that R_n exists.

Karl Popper (1902–1994), a renowned philosopher of science, suggested that only propositions or theories that can in principle be falsified, or are falsifiable, but have not actually been falsified are scientific. In other words, a necessary condition for scientific knowledge lies in the possible existence of any instances in which a corresponding proposition or theory would contradict the relevant empirical evidence and render its relation with the truth no longer valid. Duverger’s law is one of those that can potentially be falsified (e.g., R_n). On the other hand, unfalsifiable propositions and theories are not scientific but pseudoscientific, such as propaganda, fairy tales, ideology, religion, and witchcraft, which are generally imprecise and could never be proven to be false. For example, the normative proposition that proportional representation should be adopted does not represent a body of scientific knowledge since there is no way that the proposition can be falsified by denying its possible correspondence relation with the truth (if any). On Popper’s account, it makes no sense to examine whether a particular proposition or theory is true because its correspondence relation with the truth can never be verified through observational testing—that is, the generalizations obtained by inductive reasoning can at most be provisionally true until the moment when further unobserved but contradictory evidence arises in the future. Therefore, the aim of inductivism in attempting to prove any scientific proposition or theory to be true, and hence scientific knowledge, is simply misguided. Instead, according to falsificationism, the logic should be reversed such that as long as a proposition or theory is proven to be false it is not scientific (knowledge).

Popper’s distinction between science and pseudoscience is influential not only in the natural but also in the social and political sciences. The question for political scientists is this: To what extent can a proposition or theory of the political be regarded as scientific (knowledge) in accordance with falsificationism? While inductivists (and behavioralists) insist that a scientific proposition or theory must correspond to the objective social facts (the truth), falsificationists do not reject the realist standpoint concerning the truth but rather the existence of its correspondence relation with the corresponding proposition or theory being provable through

observational testing. Therefore, contrary to inductivism, falsificationism does not demand that a proposition or theory be true, but only that it not be false (and can in principle be false) in order to be scientific. Falsificationists disagree with inductivists about the role of truth in science, and Popper (1969) further posits that science progresses by conjectures and refutations instead of inductive inference from empirical evidence. In other words, a scientific proposition or theory is first constructed (conjectured) by discovering new ideas through “creative intuition,” and it is then subject to genuine tests that constitute as many difficult experiments with as much difficult evidence as possible. Such experiments and evidence are regarded as difficult if they can, at least in principle, falsify (refute) the proposed proposition or theory. If after numerous attempts at such genuine testing, the difficult evidence and experiments still do not falsify the proposed proposition or theory, then the latter is said to be corroborated, instead of proven to be true, by the former.

Falsificationism is in line with deductive reasoning. The proposed proposition or theory forms a hypothesis through conjectures for the genuine test. Suppose it is suggested that all countries adopting proportional representation exhibit multiparty systems, and such a hypothesis acts as the first premise. After running several genuine tests, it is then found that proportional representation and a two-party system are present in country A, where such evidence counts as the second premise. Therefore, the hypothesis (first premise) is said to be falsified by the evidence (second premise), and such a conclusion is reached by deduction. The influence of this Popperian view of science can be seen in the emergence of the rational choice theoretic approach in political science in the 1960s and 1970s. Unlike behavioralism, rational choice theory is based on several *a priori* assumptions as well as deductive logic in constructing propositions or theories that explain and predict the behavior of social aggregates. Similar to falsificationists, rational choice theorists do not inductively infer a scientific proposition or theory merely from empirical evidence but rather conjecture one on the foundation of some theoretical assumptions in the first instance, before the proposition or theory is subject to empirical testing. One key *a priori* assumption is that each individual is rational and behaves in a way that maximizes his or her expected utility. This represents the creative insights in constructing a scientific proposition or theory, where empirical evidence only takes up the role of genuine testing. In general, many rational choice theoretic models successfully explain some essential political phenomenon, such as the paradox of rational individual behavior but irrational collective outcomes as well as the failing of collective welfare delivery by self-interested public officials. Major seminal works include Downs (1957), Buchanan and Tullock (1962), Riker (1962), and Olson (1965).

One can immediately think of a list of criticisms of the rational choice theoretic approach to political science. In particular, the *a priori* assumptions on which any proposition and theory is based are not empirically supported, and

the general account of rationality is restrictive in the sense that it fails to capture the complexity of human behavior, which is more than atomistic and mechanistic—it is instead situated and significantly shaped by contexts, ranging from the political and social institutions to human psychology and history. That means that even if the proposition or theory is consistent with some relevant empirical evidence, it offers only an overly simplistic account of the political while other perspectives ruled out by the assumptions are left untouched. On the other hand, the lack of fit between the proposed proposition or theory and the empirical evidence does not necessarily mean that all bodies of knowledge it represents are false (which renders a complete dismissal). The mismatch may be, for example, simply due to the problems of certain *a priori* assumptions that fail to apply in every circumstance, while the proposed proposition or theory still has value for being retained. It is apparent that such a rational choice theoretic (or falsificationist) view of science in social and political inquiries may still leave something to be desired.

Departure: Scientific Paradigms and Interpretive Political Science

Inductivism and falsificationism are two scientific paradigms, which give rise to the schools of behavioralism and rational choice theory in political science respectively. They presuppose different fundamental views on whether scientific proposition or theory about (or knowledge of) the political can be obtained by confirming the correspondence relation between the truth and the empirical evidence. On the other hand, they can both be grouped under a common heading in terms of methodology. Similar to the natural sciences, both the behavioralist and the rational choice theoretic approaches aim at explaining empirical (social and political) phenomena through the construction of hypotheses and generalizations that are applicable to more than a single set of empirical observations and making falsifiable predictions about other possible observations. Given that both approaches are founded on positivism in which science and its method of acquirement are considered objective and rational, what accounts for the emergence of two paradigms that differ only along the dimension of logical reasoning (i.e., induction and deduction)? Their fundamental disagreement concerns the possibility of confirming whether a proposition or theory corresponds to the truth and whether such confirmation is necessary for the proposition or theory to be scientific. Which of these two views of science is (more) accurate, if either is at all?

Thomas Kuhn (1922–1996), a historian and philosopher of science, would focus on the “if either” part of the question above. According to Kuhn, the inductivist and the falsificationist accounts of science are misguided since neither of them adequately matches the historical evidence concerning the development of science. Science is not about the logic of justification based on observable evidence but is composed of the unobservable, which gives

rise to a particular paradigm. In other words, neither confirmation nor unfalsification of the truth suffices to determine a proposition or theory to be scientific or not. Many of the leading (natural) scientific theories were faced with certain empirical anomalies from the beginning, such as Newton's gravitational theory, but these theories were preserved and developed despite the fact that they could have been falsified at birth if the falsification demand had been strictly followed. It is also apparent that there were always adherents of established propositions and theories even if the latter were challenged by recalcitrant evidence. Kuhn reminded us that the history of science was in fact characterized by extended periods of normal science where a scientific community conducts research on a subject within an established paradigm. Such a paradigm comprises a set of theoretical assumptions that reflect the fundamental agreement among all members of the community and those members solving problems within a specific set of subjects on the basis of those theoretical assumptions. In Kuhn's view, normal science is a problem- or puzzle-solving activity, and hence, scientists dedicate their minds and time to research concerning details within an established paradigm without questioning the corresponding fundamental assumptions. For any proposition or theory that contradicts the observable evidence (anomalies), instead of being falsified right away, it is treated as simply one of the unsolved puzzles. When these anomalies are minor and limited in number, they tend to be ignored; only when they accumulate to an extent that most scientists within the paradigm begin to cast doubt on the fundamental assumptions will there be a shift of allegiance from the original paradigm to a new one, or a scientific revolution. In sum, Kuhn sees science as not only logic and rationality but also a carrier of sociological and historical implications in the form of paradigms and their progress and transitions. What makes a subject and a method to the subject scientific depends much on the consensus of the scientific community at a particular place and time. Therefore, inductivism and falsificationism are not the science as a whole but only its constituents, while these constituents (or paradigms) are mutually incommensurable.

Kuhn's sociological insights on the progress of science offer an angle of reflecting on what counts as science (knowledge) and scientific methods in social and political science. Since scientific knowledge is constructed within a paradigm by its own methods, tools, and criteria of assessment, these unquestioned rules of the game not only affect the selection of research topics but also emphasize how the findings are interpreted. This gives rise to the interpretivist tradition in social and political science. In contrast to behavioralism and rational choice theory, interpretivists insist that the goal of social and political science is to understand, rather than to generalize, empirical phenomena by conducting a rich textual, contextual, or historical (or any combination of the three) study into these phenomena. Interpretivists accept the starting point of antirealists that posits there exist no objective (social) facts, and

all social and political phenomena, whether observable or not, are structured and shaped by human thought and discourses. For instance, instead of acknowledging that it is a matter of fact that there is a positive correlation between proportional representation and a multiparty system as in Duverger's law, interpretivists are skeptical about such a robust generalization and perceive it as presenting an oversimplified, if not incorrect, version of the political. Instead, the interpretivist tradition holds that the truth is not absolute but relative, varying across space and time. This is closer to the coherence theory of truth, where something counts as true if and only if it coheres with a relevant reference set of some accepted knowledge. Here, the reference set is what interpretivists stress about the context that is shaped by a variety of factors including those that are historical and sociological. This resonates with Kuhn's view on science as paradigms. Hence, for interpretivists, it is not possible to pursue (scientific) knowledge of the political without considering the meanings behind various political phenomena existing in a particular context. For example, interpretivists seek to understand individual voting behavior by revealing the formation and justification of individuals' voting preferences. Research as such often adopts qualitative methods such as interviews and focus groups. The scope of inquiry tends to be intensive and constricted, and it is hard to avoid subjectivity in interpretations. Nevertheless, interpretivists do not intend to generalize their findings, which are instead context dependent, and subjectivity itself does not threaten the value of interpretations as (scientific) knowledge. After all, truth is relative and depends on whether (and which) reference set(s) is (are) accepted in the first place, and the latter, of course, includes the way in which the findings are interpreted. Examples of interpretivist works include Oakeshott (1962), Anderson (1991), Collingwood (1993), and Risse (1999).

Synergy: A Mixed Approach to Political Science

As discussed so far, the two major approaches—the positivist and the interpretivist—seem to be incompatible with each other because of their divergence on the fundamental views about the nature of science. A worthwhile follow-up question would be whether the two contending approaches can be reconciled to make further progress of political science research by combining both of their strengths. This concerns whether science can be both rationalistic and contextual.

Imre Lakatos (1922–1974) suggested that the ideas of Popper and Kuhn can be fused in such a way that science can be understood as not only about empirical observation and logical induction or deduction but also about taking into account the social and historical context of research. On one hand, contrary to Popper, scientific propositions or theories are not to be abandoned right away by falsification when anomalies are identified in contradiction to any empirical evidence. Instead, these propositions or theories usually coexist with the anomalies for some time, during which

scientists attempt to rescue the propositions or theories from collapsing by offering further explanation or shifting their attention to other problems. Scientific progress, according to Lakatos, is made by these rescue actions instead of by falsification, and science can be regarded as a kind of human activity, as progressive research programs. On the other hand, Lakatos disagrees with Kuhn that these programs are not mutually incommensurable paradigms resulting in relativism but are open for cross-comparisons and evaluations. Each of these programs is characterized by its own *heuristic*, which means a certain set of tools, methods, and techniques of problem solving, where the heuristic's strength determines also the strength of the corresponding research program. A heuristic is strong if propositions or theories are produced so that the corresponding research program can enable individuals to infer novel predictions that lead to the discovery of new empirical facts and explain empirical phenomena at least as well as their rival or previous propositions or theories. In other words, whether a proposition or theory is scientific depends on both the contextual factor of its belonging to a particular research program and the rational factor of the explanatory and predictive power of the corresponding research program. Science, according to Lakatos, can share a double character that covers both Popper's rationalist and Kuhn's sociological and historical aspects.

If, in principle, science is composed of both rational and contextual elements as mentioned, how may the methods of inquiry into science be reconceptualized to determine what counts as science (or knowledge) in the study of politics? King, Keohane, and Verba (1994) asserted that social and political science can be approached in both the positivist and the interpretivist manners, with the use of quantitative and qualitative research methods respectively. This is due to their shared underlying logic of interference. Propositions or theories generated from research through either method are considered scientific as long as they satisfy a set of criteria. First, these propositions or theories are meant to draw, based on empirical observations, either descriptive inferences that concern further observations or explanatory inferences that portray (causal) correlations between empirical phenomena. Here, the interpretivists would focus more on the former where the understanding of the subject matter (political phenomena) goes beyond what is observed, while the positivists would stress the latter where the (causal) correlations as such are seen as existing facts and can be revealed. Second, no matter which methods are adopted, the set of procedures for arriving at the propositions or theories should be publicly transparent and thus able to be comprehended and tested by the scholarly community. This to some extent echoes Lakatos's view that scientific research programs are mutually commensurable and open for scrutiny. Third, since there is no single perfect and absolute way of drawing inferences, any propositions or theories so generated could in principle be wrong, and hence, good scientific research should identify its degree of uncertainty. This makes the

inductivist and the interpretivist methods remain scientific even in the face of the falsificationist challenge by pardoning the demand of verification. Last, scientific research should stress both the method and the content, and the propositions or theories should be drawn by following a set of rules of inference. This is, again, compatible with both the positivist and the interpretivist positions.

A crucial implication of the previous assertions is that science is about rationality and context (Lakatos), and the positivist and the interpretivist approaches can both be applied as methods in revealing the rational and the contextual elements of science once they fulfill certain criteria (King, Keohane, & Verba). This inspires the development of political science research in the sense that methodology is no longer confined to only one approach. Instead, it is a noticeable trend that a mixed approach has become more popular since the 1980s, when the school of new institutionalism arose. The new institutionalism, which recognizes the importance of political institutions in shaping political phenomena, corresponds to Lakatos's view in the way that scientific study of the political is about not only the political phenomena per se (positivist tradition) but also the social and historical context under which they are positioned (interpretivist tradition). Meanwhile, in acquiring the science as such, the scientific method used is a combination of both quantitative (positivist tradition) and qualitative (interpretivist tradition). Examples include historical institutionalism (e.g., Pierson, 1996; Pierson & Skocpol, 2002) and rational choice institutionalism (e.g., Laver, 1997; Shepsle, 1989). Although on the operational level a synergy of the two *prima facie* contending methodologies in political science research has been brought about by the new institutionalist turn, it is not clear on the metaphysical level whether there can truly be a reconciliation between the positivists and the interpretivists concerning their fundamental disagreement about the relationship between (political) science and the truth.

Conclusion

This chapter reviews the development of the idea of science in political science. Since the 1950s, when political science arose as a separate academic discipline, there have been plenty of debates about the nature of political science as well as the corresponding method of inquiries. Although within the positivist tradition the behavioralists and the rational choice theorists disagree on the logical reasoning in arriving at propositions or theories that explain and generalize objectively existing political phenomena, the interpretivists insist on the importance of understanding the contextual factors that shape the political phenomena. Such disputes can be mapped into the disputes in the philosophy of natural and social sciences among the inductivists, the falsificationists, and the interpretivists. Although the tension can be reconciled by reconceptualizing science as comprising both rational and contextual elements and applying a mixed

approach consisting of quantitative and qualitative methods, such methodological middle ground is far from settling the conflict between the positivists and the interpretivists on the fundamental nature of (political) science.

Notes

1. Strictly speaking, political science as an academic discipline began as early as the late 1890s when the scientific study of political institutions emerged, often regarded as the old institutionalism (e.g., Spencer, 1891; Weber, 1946; and Pareto, 1935).

2. The study of macrolevel behaviors is also known as *structural functionalism*.

References and Further Readings

- Anderson, B. (1991). *Imagined communities: Reflections on the origin and spread of nationalism* (Rev. ed.). London: Verso.
- Box Steffensmeier, J., Brady, H., & Collier, D. (Eds.). (2008). *The Oxford handbook of political methodology*. Oxford, UK: Oxford University Press.
- Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent*. Ann Arbor: University of Michigan Press.
- Collingwood, R. (1993). *The idea of history* (Rev. ed.). Oxford, UK: Oxford University Press.
- Dahl, R. A. (1961). *Who governs? Democracy and power in an American city*. New Haven, CT: Yale University Press.
- David, M. (2008, Fall). The correspondence theory of truth. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved May 1, 2009, from http://plato.stanford.edu/archives/fall2008/entries/truth_correspondence
- Downs, A. (1957). *An economic theory of democracy*. London: HarperCollins.
- Duverger, M. (1964). *The idea of politics: The uses of power in society*. London: Methuen.
- Duverger, M. (1972). *Party politics and pressure groups*. New York: Thomas Y. Crowell.
- Hay, C. (2002). *Political analysis*. Basingstoke, UK: Palgrave.
- Hume, D. (1963). *An enquiry concerning human understanding*. La Salle, IL: Open Court.
- King, G., Keohane, R., & Verba, S. (1994). *Designing social inquiry*. Princeton, NJ: Princeton University Press.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Lakatos, I. (1968). Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society*, 69, 149–186.
- Laver, M. (1997). *Private desires, political action*. London: Sage.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review*, 65, 682–693.
- Lipset, S. M. (1959). *Political man*. New York: Doubleday.
- Marsh, D., & Stoker, G. (2002). *Theory and methods in political science*. Basingstoke, UK: Palgrave.
- Martin, M., & McIntyre, L. (Eds.). (1994). *Readings in the philosophy of social science*. Cambridge: MIT Press.
- Oakeshott, M. (1962). Political education. In *Rationalism in politics and other essays* (pp. 43–69). Oxford, UK: Oxford University Press.
- Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Oxford English dictionary: <http://www.oed.com>
- Pareto, V. (1935). *The mind and society*. London: Jonathan Cape.
- Parsons, T. (1966). *Societies: Evolutionary and comparative perspectives*. London: Prentice Hall.
- Pierson, P. (1996). The path to European integration: A historical institutionalist analysis. *Comparative Political Studies*, 23, 122–163.
- Pierson, P., & Skocpol, T. (2002). Historical institutionalism in contemporary political science. In I. Katznelson & H. V. Milner (Eds.), *Political science: The state of the discipline* (pp. 693–721). New York: W. W. Norton.
- Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Popper, K. (1969). *Conjectures and refutations*. London: Routledge & Kegan Paul.
- Putnam, H. (1975). *Mathematics, matter and method: Philosophical papers* (Vol. 1). Cambridge, UK: Cambridge University Press.
- Putnam, H. (1981). *Reason, truth and history*. Cambridge, UK: Cambridge University Press.
- Riker, W. H. (1962). *The theory of political coalitions*. New Haven, CT: Yale University Press.
- Riker, W. H. (1986). Duverger's law revisited. In B. Grofman & A. Lijphart (Eds.), *Electoral laws and their political consequences* (pp. 19–42). New York: Agathon.
- Risse, T. (1999). International norms and domestic change: Arguing and strategic adaptation in the human rights area. *Politics and Society*, 27(4), 526–556.
- Shepsle, K. (1989). Studying institutions: Some lessons from the rational choice approach. *Journal of Theoretical Politics*, 1(2), 131–147.
- Spencer, H. (1891). *Essays: Scientific, political and speculative*. London: Williams & Norgate.
- Truman, D. (1951). *The process of government*. New York: Knopf Press.
- Weber, M. (1946). Politics as a vocation. In H. Gerth & C. W. Mills (Eds. & Trans.), *From Max Weber: Essays in sociology* (pp. 78–128). New York: Oxford University Press.
- Young, J. O. (1995). *Global anti realism*. Aldershot, UK: Avebury.
- Young, J. O. (2008, September). The coherence theory of truth. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved May 1, 2009, from http://plato.stanford.edu/archives/fall2008/entries/truth_coherence

55

POSITIVISM AND ITS CRITIQUE

WILLIAM ROSE

Albion College

At the beginning of her 2003 presidential address to the American Political Science Association (APSA), Theda Skocpol (2004) looked back over the previous century to the New Orleans meeting in late 1903 when the association was founded. Skocpol reminded her audience that APSA was born from a conscious decision by “scholars and men of affairs” to develop an association, separate from the already-existing American Economics Association and the American Historical Association, to foster the following:

the establishment of some representative body that can take the scientific lead in all matters of political interest, encouraging research, aiding if possible in the collection and publication of valuable material and . . . in general advancing *the scientific study of politics* in the United States. (p. 1, italics added)

The aim of those present at that 1903 meeting in New Orleans was to disentangle the study of politics from history and moral philosophy and thereby to be able to demonstrate the importance of a properly scientific study of political phenomena. The political world was deemed every bit as suitable for scientific study as those realms of the natural world laid claim to by physicists and biologists. Indeed, although seemingly obvious, it is no less important to observe that of all the social sciences, it was only the discipline of political science that would choose to include

the word *science* in its name. To paraphrase Stanley Fish (1993), the discipline of political science has long wished to have a scientific existence.

The impulse for the scientific study of politics was not new of course; one might recall, for example, that “the idea that politics is, or can be, the subject of science is an ancient one that reaches back to Aristotle” (Farr, 2003b, p. 307). And the call for a new political science possessed of the capacity to meet the challenges of understanding a new world can be found in *The Federalist* as well as in Tocqueville’s study of *Democracy in America*. What was new was the effort to professionalize the study of politics—an effort greatly aided by the growth of the American university in the early part of the 20th century.

Frank J. Goodnow (1905) delivered the first presidential address to that New Orleans audience in 1903. The text of that address offers evidence of Goodnow’s resistance to the felt temptation to provide a definition of the term *political science*; rather, he was interested only to identify the proper object of its study—the state. According to Goodnow,

It may perhaps be safely said that there was not, until the formation of the American Political Science Association, any association in this country which endeavored to assemble on a common ground those persons whose main interests were connected with the scientific study of the organization and functions of the state. (p. 36)

With the scope of inquiry narrowly identified, the proper conception of science was pulled into view as well; such a conception had, for Goodnow,

to do with the State at rest and with the State in action. . . . The State, as an object of scientific study, will be considered from the point of view of the various operations necessary to the realization of the State will. (p. 37)

By contemporary standards, Goodnow's characterization seems excessively cramped, attuned primarily to legalistic and historical questions. Yet the simultaneous concern with both the scientific legitimacy of knowledge claims (here, about the state) and the desire to produce knowledge relevant to political actors still has strong resonance today.

However, not too long after Goodnow's address, some in the association would seek to reorient its focus away from a narrow emphasis on the state and concomitant questions of sovereignty, turning instead to the study of actual political processes and, more important, to the analysis of the political behavior of individual actors. Such a shift was apparent already by the sixth annual meeting of the association. In his presidential address to this gathering, A. Lawrence Lowell (1910)—recently appointed as president of Harvard University as well—suggested a new approach to the study and teaching of political science. In his address, titled "The Physiology of Politics," Lowell displayed somewhat less of the disciplinary confidence present in Goodnow's 1903 address. Rather, Lowell's presentation seemed to operate on two separate, yet related, levels: (1) an expression of disciplinary envy for advances made in the natural sciences and (2) something of a call to arms for political science to bring itself alongside the natural sciences and the more advanced social sciences (such as economics), so that political science might join in the scientific progress being made in those other disciplines. According to Lowell, too many of his professional colleagues had their gazes fixed on the clouds when, instead, they needed to anchor themselves firmly to the hard ground of fact. Their discussions were, in Lowell's estimation, too "theoretical, treating mainly of what ought to happen, rather than what actually occurs" (p. 3). Most important, for Lowell, the emphasis on abstract normative questions, at the expense of an empirically grounded understanding of actually existing political processes, rendered the discipline's knowledge claims vulnerable to the charge of irrelevance. Striving for disciplinary relevance in that area seemingly most in need of the sort of knowledge political science could promise—political reform efforts—Lowell observed,

All reform movements need for criticism a devil's advocate, who is not, however, believed to be in league with the devil; or rather they need advice from people who are really familiar with the actual working of many political institutions. In short, they need men with a scientific knowledge of the physiology of politics. (p. 3)

For Lowell, the substance of such a physiology was not to be found in books and institutions alone; rather, the realities of political life largely were to be found in the political behavior of ordinary people—behavior that could be observed scientifically with all the attendant gestures to objectivity and neutrality that a truly scientific inquiry required. Political science for Lowell was, then, the study of political behavior in all its many forms and locations.

Problems of understanding and explaining social and political phenomena certainly have manifested themselves in different forms over the years, but there are important continuities as well. Thus, a century after Goodnow's speech at Tulane University, Theda Skocpol (2004) would observe,

Despite the many changes over the decades, there have been important continuities—so many that Goodnow and his colleagues would surely recognize us today as inheritors of the association and disciplinary vision they launched. Now, as then, organized political science encompasses normative theory as well as empirical research. . . . Now, as then, we aim to link responsible citizenship in the larger society to scholarly studies of government and politics. (p. 1)

Skocpol concluded her point by quoting approvingly from yet another APSA presidential address, delivered at mid-century: "As we . . . develop political science as a discipline we both serve our professional needs and perform the vital function of helping our democracy to know itself better" (Herring, 1953, as cited in Skocpol, 2004, p. 1). The Goodnow and Skocpol presidential addresses serve as bookends to a century of APSA activity that subsumed a range of scholarly production, much of it informed (explicitly or implicitly) by the tension brought about from the desire for the enhanced legitimacy of knowledge claims made about political phenomena, and a corresponding concern that such knowledge be relevant to the real world of politics and its actors. The aspiration then, as now, was for a scientific source of knowledge about the nature, functions, and effects of political phenomena, as well the aspiration that such knowledge would have purchase in the world, that it be relevant. At times, the relationship between scientific legitimacy and relevance has manifested itself as more than a productive tension, erupting into deeply contested engagements over the very meaning of political science itself. Throughout, however, these engagements have been informed by, or played out against a backdrop provided by, positivism and the critical responses to it.

Positivism and the Aspiration for Scientific Legitimacy

What is positivism and why has it mattered to political science? How has it informed the dynamic between the legitimacy of knowledge claims about political phenomena

and the relevance of those knowledge claims to politics? As one student of the discipline has recently observed,

American political science has long aspired to emulate both the objective research methods of the natural sciences and their practical successes in controlling their objects of study. The profession's mainstream aspires to establish a cumulative, stable body of objective knowledge about politics at the same time that many political scientists wish to use such scientific knowledge to shape politics, for example by advising policy makers or contributing to public debates. (Oren, 2006, p. 73)

The effort to model social and political inquiry on the natural sciences is at the heart of the positivist impulse.

Pinning down just what one is talking about when discussing positivism—generally, or in the context of the discipline of political science—is, however, a much more complicated task than at first it might seem. Beyond the fact that there are several different traditions and streams of positivist thought, it is also unclear whether (or in what form) positivism continues to exist. For some, it is still the case that the appellation of the term *positivist* is, as James Johnson (2006) said, a “badge of honor, worn . . . to identify those whose research is seen—if not actually, then at least potentially—as embodying the virtues of rigor, clarity, and solidity” (p. 225). For others, the label is one to be avoided. For example, according to Anthony Giddens (1977), positivism “has today become more a term of abuse than a technical term of philosophy” (p. 29). Undoubtedly, at least in recent years, the term has been deployed as “a sufficient reason to dismiss entire brands of research and those who conduct them as abstract, sterile, and politically dogmatic in disciplinary and extradisciplinary terms” (Johnson, 2006, p. 225). Nonetheless, the concept seems to retain a central, if somewhat ambiguous, role within the social sciences generally and political science more particularly. It is perpetually disavowed yet often unconsciously embraced as a default orientation to ground scientific research in the social sciences. Positivism has been declared an anachronism at various points throughout the 20th century, only to reemerge with an uncanny persistence (see, e.g., Steinmetz, 2005a). From the perspective of the philosophy of social science, Richard Miller (1987) observed that “few people, these days, call themselves ‘positivists.’ In fact [according to Miller], I have never met a self-proclaimed positivist. Yet in a broad sense of the term, positivism remains the dominant philosophy of science” (p. 3). A somewhat more troubled image has been offered by George Steinmetz (2005a), who characterizes positivism as a “specter” that continues to haunt the social sciences—a spectral presence that endures “despite repeated attempts by social theorists and researchers to drive a stake through the heart of the vampire” (p. 3).

Further complications emerge when one confronts the fact that there are multiple positivisms. That is, it is fair to say that positivism has at least four meanings. First, one can associate the term with Auguste Comte's theory of social

evolution, an approach to the comprehensive understanding of the social world, an understanding greatly expanded and conceptually developed in the social theory of Émile Durkheim. Or the term can be associated with the philosophical tradition of thought from the early part of the 20th century known as *logical positivism*. Further still, the term is identified with a tradition in Anglo-American jurisprudential thought known as legal positivism, a philosophy of law that finds its most sophisticated expression in the work of the late H. L. A. Hart (see, e.g., Hart, 1997). Finally, the term might be taken to refer to a set of research practices associated with (social) scientific inquiry—what has been characterized elsewhere as *methodological positivism* (see, generally, Steinmetz, 2005b). Although there are conceptual linkages and historical relationships that can be drawn among these four iterations of positivism, this chapter's principal concern is with this last meaning of the term: an understanding of positivism that “refers to a concept of knowledge, a concept of social reality, and a concept of science” (Riley, 2007, p. 115).

Such complicating factors aside, it is still possible to bring into focus some central features of the continuing positivist orientation in American political science. The positivist temper still dominant among many political scientists may be characterized as one that conceives of political inquiry as a form of scientific inquiry “that differs in degree and not in kind from the well-established natural sciences,” and these same political scientists are “convinced that the greatest success is to be found in emulating, modifying, and adapting techniques that have proven successful in our scientific understanding of nature” (Bernstein, 1976, p. xv). Certainly, there are differences within the discipline as to the “essential features” of a positivist political science (Meilleur, 2005), but Timothy Kaufman-Osborn (1991) has nonetheless provided a helpful definition that subsumes a range of approaches under the larger banner of positivism: Positivism is a theory of science that does the three following things:

- (1) advances a nomological conception of knowledge, one that identifies the end of inquiry with the construction of causal explanations relating the occurrence of specific events through reference to universal laws that predict an invariant relationship between antecedent conditions and their necessary consequences; (2) claims that a presupposition of such knowledge is the generation of a neutral language whose content stands in some isomorphic relationship to the antecedently existent objects it describes; and (3) affirms the ideal of value free knowledge. (p. 229)

Beyond definitions, positivism has a history that is important to note. The intellectual origins of positivism traditionally are traced to the work of Auguste Comte (1798–1857). According to Raymond Williams (1985), Comte's positivism “was not only a theory of knowledge; it was also a scheme of history and a programme of social reform” (p. 239). Throughout the 19th century,

positivism thus connoted “a free-thinking and radical as well as a scientific movement” (p. 239). Comte’s efforts at social reform were born of a desire for a more equitable society, and for Comte, a more comprehensive science of society was seen as the path to that desired end. Comte’s conception of a positive science was born, in other words, from normative or ethical concerns. That is, Comte sought to import and apply the approach and methods of the natural sciences to the study of society with the aim of improving the human condition. In this, Comte continued down a path first gestured toward by his mentor Henri Saint-Simon (1760–1825). Expressing his dissatisfaction with the highly subjective nature of political inquiry and the aspiration that it be made over in the image of the natural sciences, Saint-Simon (anticipating the scientific aspirations of A. Lawrence Lowell) observed,

Hitherto, the method of the sciences of observation has not been introduced into political questions; every man has imported his point of view, method of reasoning and judging, and hence there is not yet any precision in the answers, or universality in the results. The time has come when this infancy of the sciences should cease. (Thiele, 1997, p. 9)

Thus, what would come to be called positivism emerged from Saint-Simon and Comte’s desire to purge scientific inquiry of the excess baggage of metaphysics. For Comte especially, positivism “came to signify the nineteenth-century desire to make natural science the sole model of knowledge, even for inquiries into human history and culture” (Lane, 2003, p. 321).

Comte’s (1988) positivism was grounded in an all-encompassing view of the nature of human and social evolution. According to Comte, historical observation revealed that human society had evolved through three principal stages of history. As put forward in his *Introduction to Positive Philosophy*, Comte explains that both the human mind and society have exhibited a progressive growth. As such, said Comte,

In thus studying the total development of human intelligence in its different spheres of activity, from its first and simplest beginning up to our own time, I believe that I have discovered a great fundamental law, to which the mind is subjected by an invariable necessity. . . . This law consists in the fact that each of our principal conceptions, each branch of our knowledge, passes in succession through three different theoretical states: the theological or fictitious state, the metaphysical or abstract, and the scientific or positive state. . . . In the theological state, the human mind directs its researches mainly toward the inner nature of beings . . . in a word, toward absolute knowledge. . . . In the metaphysical state . . . the supernatural agents are replaced by abstract forces, real entities or personified abstractions, inherent in the different beings of the world. . . . Finally, in the positive state, the human mind, recognizing the impossibility of obtaining absolute truth, gives up the search after the origin and hidden causes of the universe

and a knowledge of the final causes of phenomena. It endeavours now only to discover, by a well combined use of reasoning and observation, the actual laws of phenomena. (pp. 1–2)

Although highly influential in mid-19th-century European thought, Comte’s specific influence would be on the wane already by the end of the century. Yet his desire for a scientific understanding of society, shorn of the distorting influence of metaphysical abstraction, would remain. And indeed, the scientific aspiration for the study of political life expressed by Goodnow and Lowell suggests a general family resemblance with Comte. For Goodnow and Lowell, however, the orientation could be characterized more generally as a sort of methodological positivism. Envious of the growing success—and influence—of the natural sciences, their efforts might be articulated in terms of a desire to understand and explain social and political phenomena from a position informed by and compatible with the natural sciences—objective, reliant on reason, grounded in empirical observation, and value neutral. In the broadest sense, this positivism was grounded, Miller (1987) wrote,

in the assumption that the most important methodological notions for example, explanation, confirmation and the identification of one kind of entity with another can each be applied according to rules that are the same for all sciences and historical periods, that are valid a priori, and that only require knowledge of the internal content of the propositions involved for their effective application. (p. 3)

What, then, defines the positivist position that this chapter is concerned with, what has been identified as methodological positivism? First, since its emergence with Comte, most of those who have been committed to a positivist orientation also have been largely committed to the notion of the unity of scientific inquiry. Such a position is grounded in a series of related assumptions, principally “that the universe is a causally ordered, homogenous, one-layer world, that there is a basic unity to human experience and that we are therefore able to gain knowledge of reality and indeed construct a knowledge system about it” (Delanty & Strydom, 2003, pp. 13–14). Entailed by this is the further claim that it is therefore possible—indeed desirable—to produce a scientific discourse subsuming all areas of scientific inquiry, both natural and social. Distinctions between the natural and social sciences would dissolve, and a unified scientific methodology and language would emerge.

Further, positivism has often been articulated in a form that links a discernible concept of social reality (an ontology) with a corollary self-understanding of scientific inquiry. Specifically, positivism is informed by an ontology “that equates existence with objects that are *observable*,” and it is associated with an approach to social science wherein scientific inquiry is understood to be “independent of the reality it describes” (Riley, 2007, p. 115, italics added). This

antimetaphysical orientation thus cautions against mistaking the abstract categories of thought for the concrete real. Instead, positivists stress that the basis of all knowledge claims must be grounded in the positive data of experience. Abstract speculation about the nature of the world or normative assertions as to how the world should be, divorced from concrete evidence supplied by observations of the world as experienced, could not be the basis of valid knowledge claims about the world. This emphasis on the experiential basis of all valid knowledge claims has manifested itself in two related ways:

either phenomenalism, in which case the emphasis is on the immediate experience of phenomenal or mental entities in the form of observables or sense data; or physicalism (or naturalism), in which case the emphasis is on perceptual or physical entities or common sense things and events that can be intersubjectively verified by recourse to empirical evidence. (Delanty & Strydom, 2003, p. 15)

Accordingly, the empirical realm—composed of the data that can be observed and confirmed by the senses (e.g., observation)—exhausts the field of legitimate scientific investigation, the field about which legitimate knowledge claims can be made. Therefore, “only those knowledge claims founded directly upon observable experience can be genuine” (Hawkesworth, 2003, p. 8).

Two related claims historically have been intertwined within positivism. First, traditionally, it has rested on an assumed separation between the knower and the known, a distinction drawn between the subject and object of scientific knowledge. That is, in its various iterations, positivism presumes that scientific inquiry begins with neutral observation. Moreover, the validity of truth claims about the observed world is bound up in this putative separation between the observer and the observed. The facts are said to speak for themselves, and as such, knowledge claims about the facts as observed are not contingent on the idiosyncrasies of the observer but linked instead to the methods used to ascertain and explain those facts. To put this another way, positivism assumes that facts are given and that observation of the raw data of the world is unmediated. Any scientific explanation, therefore, may unhesitatingly focus on the object of inquiry alone. This entails a second, yet related, separation—that between fact and value. That is, not only is an unmediated apprehension of the facts that form the object of social scientific inquiry possible, but it also is the case that the normative valuations of the observer must not intrude on the process of inquiry:

Based on the assumption of the necessity of upholding a logical separation of facts and values or descriptive and normative statements, the demand is made that [social] science should proceed in a neutral manner, free from all infection by personal, ethical, moral, social or cultural values, with the scientist actively desisting from deriving “ought from is” or “values from facts.” (Delanty & Strydom, 2003, p. 14)

As others have observed, positivism “is above all a position within epistemology,” in that it insists that “scientific explanations take the general form of ‘if A then B’” (Steinmetz, 2005a, p. 32). In other words, positivism identifies valid scientific knowledge with certain covering laws. One can distinguish valid knowledge claims from a simple assertion of opinion if there exists some means by which to confirm the truth of the claim. Casting the knowledge claim in an if-then form provides the means by which the validity of the claim can be tested. Descriptions of observed facts, as those facts are subsumed within the if-then causal relationship, yield theories about these relationships, from which hypotheses may be deduced. Importantly, for positivism, valid scientific theory is the result of the observation and is not formed prior to the observation. From the standard positivist perspective, then, theory may be understood as composed of interrelated concepts and propositions joined in order to provide systematic claims about social phenomena, with the aim of understanding, explaining, and predicting the phenomena. The emphasis on the capacity of positivist social science to make predictions about the world leads to the final attribute of positivism that should be taken into account. Specifically, positivist social science often seeks not just to comprehend the social and political world but to control or manipulate its objects of study as well (see, e.g., Oren, 2006), thus combining the concern to produce legitimate knowledge claims about politics with the concern that such knowledge claims be relevant to politics.

It is fair to say that political scientists have differed over the central elements of what might constitute a positivist approach to social and political inquiry. Common themes are nonetheless discernible: The goal of political inquiry is the construction of causal explanations “relating the occurrence of specific events through reference to universal laws that predict an invariant relationship between certain antecedent conditions and their necessary consequences” (Kaufman-Osborn, 1991, p. 229); the presumption of neutrality in observation and description of those events; and the affirmation of the scientific ideal of value-free inquiry.

Positivism, Behavioralism, and the Struggle for Theory

The high tide of positivist aspirations in American political science came after World War II with the emergence of behavioralism. Not so much a school of thought, behavioralism was more an orientation toward the relationship between knowledge and politics. There are nonetheless specific individuals and institutions that feature prominently in the history of behavioralism; for present purposes, the University of Chicago and David Easton may be considered to stand in for larger developments within the field.

As early as the 1920s, something more than the mere desire that the study of politics become scientific had been articulated by members of the discipline. A central figure in these discussions was Charles Merriam, who was located in the Department of Political Science at the University of Chicago. Impatient with the perceived inability of the discipline to speak meaningfully to contemporary political matters and to lend relevant expertise to political reform efforts, Merriam (who was instrumental in founding the Social Science Research Council in 1923) called for a political science that was to be explicitly modeled on the natural sciences. Most important, Merriam contended that this new science of politics takes political behavior as its principal object of study. In his seminal article "The Present State of the Study of Politics," Merriam (1921) made the case for moving the study of political behavior to the center stage of scientific political inquiry. He argued that from that point on, political scientists should devote their energies to the observation, measurement, quantification, theorization, and prediction of political behavior. What would come to be called the Chicago School of Political Science, under Merriam's leadership, pressed for a natural science model for the study of politics. Merriam's concern, however, was not only with the scientific legitimacy of the knowledge produced by the discipline, but also with the relevance of scientific political analysis for practical politics.

Although Merriam's work developing a science of political behavior was central to the development of a positivist political science, it was not until after World War II that positivism, as encapsulated in a behavioralist political science, would rise to prominence. And here, it was Merriam's former student, David Easton, who assumed a central role in the development of positivism within political science. Easton (1953) published *The Political System: An Inquiry Into the State of Political Science*, a text that has been dubbed "behavioralism's manifesto" (Farr, 2003a, p. 443). Easton (1953) introduced his study by making clear the stakes:

This study deals with the condition of the science of politics as it is known in the United States today and with the relation to it of general political theory. . . . In earlier centuries a student of political systems would have been hard put to distinguish his theoretical inquiries from his general political research. Today, in the United States, however, it has become increasingly difficult to appreciate why political theory should continue to be included as a central part of political science. Theory has become increasingly remote from the mainstream of political research. This study will have served a useful purpose if it helps in some small way to win back for theory its proper and necessary place. (p. ix)

Easton's self-defined task, in other words, was to save political theory from itself and restore it to a central place within the discipline. But Easton's theory would be emptied of its historicist fascination with the canon of so-called great thinkers. He would seek instead to develop a properly

positive political theory, and once "rejuvenated"—that is, empirically oriented—political theory would be centrally placed to further the development of "reliable political knowledge" (p. 309). Like earlier aspirants to the scientific study of politics, Easton too sought to enhance the legitimacy of the discipline's knowledge claims. More than ever, the appeal to scientific method seemed to hold the promise of successfully trading on the cultural authority of the natural sciences. Less concerned was Easton, however, with making claims for the direct relevance of such knowledge for political reform efforts. Initially, at least, this was social scientific inquiry for the sake of knowledge itself.

For the next decade and a half, the positivist orientation of behavioralist work in political science could be viewed as dominant within the discipline. Although the behavioralist approach itself never could be understood to describe the bulk of research in the discipline, a positivist undercurrent largely informed the dominant aspirations of the profession. And as such, positivism's prominence ultimately invited a harsh reaction. Moreover, the real world of politics intruded in ways that called for a reassessment of the nature and ends of political inquiry. What did political scientists know about politics and why did any of the knowledge about politics produced within the discipline matter to those outside the discipline?

Easton (1965) published a second key text in the development and articulation of behavioralism, *A Framework for Political Analysis*. There, he identified a "revolution in political theory" that was then under way: the emergence of an empirically oriented political theory "often referred to as behavioral theory" (p. 3). Recognizing the many versions that passed under the banner of behavioralism, he nonetheless found common ground. According to Easton, "Closer inspection does reveal that they are all looking ahead toward the same region in space—a science of politics modeled after the methodological assumptions of the natural sciences" (p. 8). For Easton, the behavioral approach was evidence of the discipline's advancement toward a new stage of scientific maturity.

What was the credo of this revolutionary movement? Famously, Easton (1965) identified eight major tenets of behavioralism: (1) existence of discoverable regularities in political behavior that can be "expressed in generalizations or theories with explanatory and predictive value"; (2) insistence on the need for verification; that is, "the validity of such generalizations must be testable"; (3) identification of the importance of developing appropriate techniques to acquire and interpret data; (4) enhancement of the place of quantification in political inquiry and developing appropriate quantitative methods; (5) recognition of the importance of maintaining the distinction between fact and value; (6) an increase in efforts to systematize political research; (7) affirmation of the logical priority of acquiring knowledge about political behavior over any attempt to apply such knowledge to the problems of society; and (8) recognition of the importance of the integration of the social sciences (p. 7).

Yet before the decade was out, Easton's credo would be effectively abandoned (at least by him), and this first revolution was declared to be at an end. Did this mean that behaviorist approaches to political studies were to be jettisoned or that the positivist orientation in political inquiry was rejected? No. However, strong criticism from within the discipline, combined with the pressure of external political events, brought smoldering tensions to light and forced a very public reckoning of the nature of political inquiry and the meaning and relevance of the knowledge it produced. Easton (1969) began his 1968 presidential address to APSA by observing,

A new revolution is under way in American political science. The last revolution—behavioralism—has scarcely been completed before it has been overtaken by the increasing social and political crises of our time. The weight of these crises is being felt within our discipline in the form of a new conflict in the throes of which we now find ourselves. This new and latest challenge is directed against a developing behavioral orthodoxy. This challenge I shall call the post behavioral revolution. (p. 1051)

In his presidential address, Easton (1969) acknowledged the “deep dissatisfaction with political research and teaching” (p. 1051) that was being voiced from various corners of the discipline. Most important, the discipline's aspirations to scientific legitimacy and political relevancy were under direct assault. In the wake of the civil rights protests of the 1950s and 1960s, the riots in several of America's major cities in the mid- to late 1960s, and the escalation of the Vietnam War with its attendant social unrest back home, those most closely associated with a positivist orientation were forced to reconsider “the very possibility of a science of politics” (p. 1051). For many, behavioralism seemed particularly ill equipped to understand, let alone predict, the many social and political crises afflicting the country and the world. The political knowledge produced from behaviorist research programs seemed out of touch, irrelevant to actual politics and social problems. Easton himself was concerned that the discipline risked appearing “more as apologists” than as “objective analysts” of current U.S. government policy. For Easton, then, a new revolution was needed, one whose “battle cries are *relevance* and *action*” (p. 1051). He was quick to claim, however, that the current critical response to behavioralism was different in important ways from earlier reactions to efforts at incorporating scientific methods into the study of politics. According to Easton,

Hitherto resistance to the incorporation of scientific method has come in the form of an appeal to the past—to classical political science, such as natural law, or to the more loosely conceived non methodology of traditional research. Behavioralism was viewed as a threat to the status quo; classicism and traditionalism were responses calculated to preserve some part of what had been, by denying the very possibility of a science of politics. (p. 1051)

As Easton (1969) represented matters, a defining characteristic of the post-behavioral revolution was that it resisted looking back to some golden age of political inquiry; instead, this new revolution was informed by work that was relentlessly future oriented: “a genuine revolution, not a reaction, a becoming, not a preservation” (p. 1051). The essential nature of this work was captured in what he called its “credo of relevance,” harkening back to the “credo of behavioralism” Easton had articulated just a few years before. Acknowledging the importance of the discipline's ability to collectively rise up to meet the press of contemporary political events, Easton nonetheless insisted that political science must not abandon its positivist orientation. “We do not need to abandon the historical objectives of basic science. There is a strategy that will enable us to respond to the abnormal urgency of the present crises and yet still preserve these traditions” (p. 1055). In the end, however, Easton's identification of a post-behavioral revolution might be understood as an attempt to acknowledge shortcomings of behavioralism—principally, its turning away from the real world of politics—while striving to reinvigorate behavioralism itself, to make the scientific study of politics relevant and responsive to the politics and policies of the moment. As Easton put it, “By adopting this course, post-behavioralism need not be considered a threat to behavioral research but only an extension of it necessary for coping with the unusual problems of the present epoch” (p. 1055).

There were, of course, those within the discipline who were deeply critical of behavioralism and who found support for this critique in the past that Easton wanted to leave behind, in the tradition of political thought. This was not merely nostalgia for the wisdom of the ancients, however, nor valorization of the past for its own sake. Rather, the effort was one of retrieval—seeking to meet the demands of the present political moment with the resources of the past of political philosophy. Easton strove to develop a positivist political theory informed by a properly scientific sensibility. Contemporaries such as Eric Voegelin and Leo Strauss (the latter located, like Easton, at the University of Chicago), on the other hand, sought to restore morality to political inquiry, to merge fact and value, and to eschew efforts aimed at aligning political inquiry with the scientific study of the natural world. In essence, this was a struggle over the very nature of political theory. But this struggle for theory concerned something even more central: It concerned how one sees the political world and what one can legitimately claim to know about that world. And it concerned the relevance of such knowledge claims to politics itself. Rather than engage with positivist arguments, however, Strauss (1962) adopted a dismissive stance. In his “Epilogue” to a critical study of positivist political science, authored by several of his former students at the University of Chicago, Strauss wrote,

Only a great fool would call the new political science (i.e., behavioralism and its fellow travelers) diabolic: it has no attributes peculiar to fallen angels. It is not even Machiavellian,

for Machiavelli's teaching was graceful, subtle, and colorful. Nor is it Neronian. Nevertheless one may say of it that it fiddles while Rome burns. It is excused by two facts: it does not know that it fiddles, and it does not know that Rome burns. (p. 327)

For Voegelin (1987), positivism—and its scientific aspirations—stood as a major obstacle to the further development of Western civilization. In *The New Science of Politics*, he situated the political science of Aristotle in critical juxtaposition to the misleading claims of positivism. What was needed, he claimed, was a restoration of political science to its true principles. Such restoration was needed because of the destructive effects of positivism itself. Beginning in the second half of the 19th century, according to Voegelin, positivism increasingly threatened to destroy true scientific knowledge. From Voegelin's perspective, the destructive work of positivism was the result of the operation of two related assumptions:

In the first place, the splendid unfolding of the natural sciences was co responsible with other factors for the assumption that the methods used in the mathematizing sciences of the external world were possessed of some inherent virtue and that all other sciences would achieve comparable success if they followed the example and accepted these methods as their model. (p. 4)

Voegelin understood this argument for the unity of science, on its own, as relatively harmless. However, "It became dangerous because it combined with the second assumption that the methods of the natural sciences were a criterion for theoretical relevance in general" (p. 4). It was this second assumption that was the real source of danger, according to Voegelin:

From the combination of the two assumptions followed the well known series of assertions that a study of reality could qualify as *scientific* only if it used the methods of the natural sciences, that problems couched in other terms were illusory problems, that in particular metaphysical questions which do not admit of answers by the methods of the sciences of phenomena should not be asked, that realms of being which are not accessible to exploration by the model methods were irrelevant. (p. 4)

Perhaps the most influential critique of positivism to emerge during this period would appear the same year as Easton's presidential address. Writing from a rather different political orientation than Voegelin, the influential political theorist Sheldon Wolin (1969) also focused (like Voegelin) on the problem of method (the emphasis in positivism on the importance of scientific method to the production of valid knowledge claims about the world) in his critique of positivist political science. But in his seminal essay, "Political Theory as a Vocation," Wolin would focus on the constitutive effects of what he characterized as methodism. According to Wolin, the positivist emphasis on method or technique limited the kinds of questions deemed important

or relevant for investigation. Wolin wanted to preserve some critical vantage point from which to engage with the urgent problems of the day; this was not a vantage point located at some Archimedean position above politics but one that was fully informed by normative considerations. For Wolin, the positivist emphasis on method unwittingly invited a form of political quiescence—or worse, complicity. What were the guiding assumptions or framework that the methodist followed?

The answer . . . is that there is such a framework of assumptions. It is the ideological paradigm reflective of the same political community which the normal scientists are investigating. Thus when a researcher takes "the normal flow of events in American politics" as his starting point, it is not surprising to find him concluding that "the long run stability of the system depends on the underlying division of party loyalties." (p. 1064)

In Wolin's (1969) view, methods are not merely neutral tools, to be picked up and used when and where appropriate. Rather, "methodism is ultimately a proposal for shaping the mind. Social scientists have sensed this when they have noted that research methods are 'tools' which can become a way of looking at the world, of judging everyday experience" (p. 1064). For Wolin, then, it was important to understand how political science sees the world; his concern, in other words, was with understanding the importance of the politics of knowledge. There is no neutral vantage point of investigation, no purely objective vision of the object of investigation. Rather than simply observing facts in the world, method participates in the production of those facts, which it then names as true. The legitimacy of knowledge claims about politics and the relevance of those knowledge claims to politics were to be understood accordingly. In Wolin's estimation, the primacy of method to the positivist orientation promoted an inherently uncritical view of the world:

The alleged neutrality of a methodist's training overlooks significant philosophical assumptions admittedly incorporated into the outlook of those who advocate scientific inquiry into politics. These assumptions are such as to reinforce an uncritical view of existing political structures and all that they imply. For the employment of method assumes, even requires, that the world be of one kind rather than another if techniques are to be effective. (Wolin 1969, p. 1064)

For Wolin, then, such an approach invites not merely political quiescence but something more: "Method is not a thing for all worlds. It presupposes a certain answer to a Kantian type of question, *What must the world be like for the methodist's knowledge to be possible?*" (p. 1064, italics added).

For Wolin, the task of political inquiry—of political theory—was an inherently critical one, aimed at providing individuals with the knowledge necessary to be able to interpret their experience of the world. Wolin's principal

concern, then, was with the relevance of political theory and what should be the relationship between theoretical inquiry and the so-called real world. Theory, for Wolin, provided a site from which the theorist could critically engage with the contemporary political world and imagine how things might be otherwise. The issue was, according to Wolin (1969),

not between theories which are normative and those which are not: nor is it between those political scientists who are theoretical and those who are not. Rather, it is between those who would restrict the “reach” of theory by dwelling on facts which are selected by what are assumed to be the functional requisites of the existing paradigm, and those who believe that because facts are richer than theories, it is the task of the theoretical imagination to restate new possibilities. (p. 1082)

Wolin viewed the work of positivists such as Easton as “deflationary” of this very possibility.

After Behavioralism

Behavioralism, as a self-identified program of political inquiry, tended to fade into the background of the discipline by the end of the 1970s. Yet the positivist impulse would remain, reemerging with approaches such as rational choice theory and in efforts to reconcile qualitative with quantitative approaches to political research (see, e.g., King, Keohane, & Verba, 1994). And what would count as systematic or scientific political inquiry would become more deeply contested and unsettled. Political science, as science, would be forced to contend with alternative viable claims to social and political description and analysis. Interpretivist approaches to understanding political phenomena, though present as early as the founding of APSA (see, e.g., Weber as quoted in Gunnell, 2006), would emerge with greater definition and influence in the 1970s and 1980s. Indeed, one might say that contemporary debates over the meaning and relevance of different approaches to political inquiry are now framed by the terms that positivism and interpretivism provide.

Foreshadowed by earlier critiques of positivism (see, e.g., Storing, 1962; Wolin, 1969), interpretivism challenges many of the fundamental assumptions of the positivist approach (see, e.g., Rabinow & Sullivan, 1987a). Specifically, interpretivism, in its most basic iterations, holds that observation is always culturally mediated in the form of the language drawn on to describe the world; in other words, language is an artifact of time and place. In addition, many interpretivists contend that theory over-determines and structures our observation of the world—that is, that scientific observation is always already theory laden, as opposed to theory following from scientific observation as positivism would have it. The interpretive approach seeks instead to focus attention on the “concrete varieties of cultural meaning, in their particularity and complex nature” (Rabinow & Sullivan, 1987b, pp. 5–6).

Accordingly, an interpretive approach begins from the assumption that the web of cultural meaning is itself constitutive of human experience, and this shared world of meaning forms the basis through which “the subjects of human discourse constitute themselves” (p. 6). The interplay of these shared meanings and their constitutive effects cannot be merely observed and described but must be interpreted.

Further, interpretivists question how positivists might categorize the objects to be studied in such a way as to ensure that the same discrete social and political phenomena are being observed by other researchers, as a positivist approach would require. Squeezing the complex phenomena of the political world into abstract categories for the purpose of description and explanation seems to risk missing much of what might be important to understanding the nature and meaning of the phenomena in question. Moreover, a scientific approach to the study of politics relies on quantification: to describe, to discriminate observed facts, and to generate empirical data amenable to statistical analysis in order to describe the phenomena, posit causal relationships, and generate theories to explain the phenomena. But interpretivists question the very idea of quantifying actions and behaviors that may be laden with a complex array of subjective meanings—meanings that are centrally important to understanding the phenomena in question but otherwise inscrutable to the quantitative measures of positivist political analysis.

Interpretivism, then, is a catchall term that subsumes a range of sometimes loosely connected approaches to social and political inquiry—their shared characteristics brought into sharper relief by identifying their mutual opposition to positivism. As such, the interpretivist approach generally signals a commitment to closing the positivist divide between normative claims and descriptions of fact, an approach born from the belief that evaluative claims and statements of fact are, in practice, virtually impossible to disentangle from one another. For interpretivists, there is no Archimedean point, outside language and outside history, from which to make such statements of fact. Rather, all knowledge claims are historically and culturally contingent, situated in time and place and, as such, artifacts of relations of power.

Conclusion

The contemporary presence and importance of positivism can be seen clearly, if somewhat more indirectly, through the emergence of a protest movement within American political science, Perestroika (see, generally, Monroe, 2005). The Perestroika movement arose in 2000, initially formed as a small group of academic political scientists and graduate students that expressed their profound impatience with “what they saw as the narrow parochialism and methodological bias toward the quantitative, behavioral, rational choice, statistical, and formal modeling approaches” (Monroe, 2005, p. 1) that continue to dominate the scholarly

output of much of American political science. On one level, the demand of the Perestroikans can be understood as a call for greater methodological pluralism within political science—creating greater space within the discipline for qualitative and interpretive approaches. On another, deeper level, however, the Perestroika movement reanimates many of the important tensions that have occupied the discipline since its inception in 1903. What can political scientists claim to know about the world, and what is the relevance of the knowledge they produce about politics to political actors and institutions? For example, the Perestroika movement resembles in many important ways the Caucus for a New Political Science that emerged in the late 1960s in response to the felt sense among those in the Caucus that the discipline's increasing infatuation with scientific method inhibited its ability to say anything meaningful about politics itself (see, e.g., McCoy & Playford, 1967). The positivist preoccupation that political science become more like the natural sciences and thereby bolster the legitimacy of its claims to knowledge was viewed by those in the Caucus as leading the discipline away from its more central role of improving the practice of politics.

This tension between the legitimacy of knowledge claims about politics and the relevance of those claims to politics is an enduring one; the specific topography of the debate has changed over the years, but the basic contours of the engagement have remained remarkably consistent—and the terms of that engagement have largely been defined by reference to positivism and its critics. In his discussion of the founding of APSA in 1903, John Gunnell (2006) quotes from an essay published the following year by the German social theorist Max Weber titled “The ‘Objectivity’ of Knowledge in Social Science and Social Policy.” Gunnell observed,

Although he [Weber] presented his essay as an intervention in controversies about the nature of social scientific explanation, he also explicitly addressed it to a wider public audience with the aim of vouchsafing the cognitive authority of academic social science. . . . Weber . . . emphasized various ways in which social scientific knowledge could, in principle, constrain and direct policy decisions as well as the extent to which scientific investigation necessarily proceeded from the perspective of value laden premises. The authority of social science nevertheless depended . . . on acceptance of the autonomy of empirical claims and on the professional . . . independence of those who made such claims. The dilemma and solution Weber articulated bore remarkable similarities to the situation attending the founding of the [APSA]. (pp. 480–481)

References and Further Readings

- Bernstein, R. J. (1976). *The restructuring of social and political theory*. New York: Harcourt Brace.
- Comte, A. (1988). *Introduction to positive philosophy* (Frederick Ferre, Ed. & Trans.). Indianapolis: Hackett.
- Crick, B. (1953). *The American science of politics: Its origins and conditions*. London: Routledge & Kegan Paul.
- Dahl, R. A. (1961). The behavioral approach in political science: Epitaph for a monument to a successful protest. *American Political Science Review*, 55, 763–772.
- Delanty, G., & Strydom, P. (Eds.). (2003). *Philosophies of social science*. London: Open University Press.
- Easton, D. (1953). *The political system: An inquiry into the state of political science*. New York: Knopf.
- Easton, D. (1965). *A framework for political analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Easton, D. (1969). The new revolution in political science. *American Political Science Review*, 63, 1051–1061.
- Farr, J. (2003a). The new science of politics. In T. Ball & R. Bellamy (Eds.), *The Cambridge history of twentieth century political thought* (pp. 431–445). Cambridge, UK: Cambridge University Press.
- Farr, J. (2003b). Political science. In T. Porter & D. Ross (Eds.), *The Cambridge history of science: Vol. 7. The modern social sciences* (pp. 307–328). Cambridge, UK: Cambridge University Press.
- Fish, S. (1993). The law wishes to have a formal existence. In A. Sarat & T. Kearns (Eds.), *The fate of law* (pp. 159–201). Ann Arbor: University of Michigan Press.
- Giddens, A. (1977). *Studies in social and political theory*. New York: Basic Books.
- Goodnow, F. (1905). The work of the American Political Science Association. In *Proceedings of the American Political Science Association: Vol. 1. First annual meeting, 1904* (pp. 35–46). Lancaster, PA: Wickersham Press.
- Gunnell, J. G. (2006). The founding of the American Political Science Association. *American Political Science Review*, 100, 479–486.
- Hart, H. L. A. (1997). *The concept of law* (2nd ed.). Oxford, UK: Oxford University Press.
- Hawkesworth, M. (2003). Political science in a new millennium: Issues of knowledge and politics. In M. Hawkesworth & M. Kogan (Eds.), *Encyclopedia of government and politics* (2nd ed., pp. 3–38). New York: Routledge.
- Herring, P. (1953). On the study of government. *American Political Science Review*, 67, 961–974.
- Johnson, J. (2006). Consequences of positivism. *Comparative Political Studies*, 39, 224–252.
- Kaufman Osborn, T. (1991). *Politics/sense/experience: A pragmatic inquiry into the promise of democracy*. Ithaca, NY: Cornell University Press.
- King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Lane, M. (2003). Positivism: Reactions and developments. In T. Ball & R. Bellamy (Eds.), *The Cambridge history of twentieth century political thought* (pp. 321–342). Cambridge, UK: Cambridge University Press.
- Lowell, A. L. (1910). The physiology of politics. *American Political Science Review*, 4, 1–15.
- McCoy, C. A., & Playford, J. (Eds.). (1967). *Apolitical politics: A critique of behavioralism*. New York: Thomas Y. Crowell.
- Meilleur, M. J. (2005). After methodology: Toward a profession of political science. In K. R. Monroe (Ed.), *Perestroika! The raucous rebellion in political science* (pp. 489–506). New Haven, CT: Yale University Press.

- Merriam, C. E. (1921). The present state of the study of politics. *American Political Science Review*, 15, 173–185.
- Miller, R. (1987). *Fact and method: Explanation, confirmation and reality in the natural and social sciences*. Princeton, NJ: Princeton University Press.
- Monroe, K. R. (Ed.). (2005). *Perestroika! The raucous rebellion in political science*. New Haven, CT: Yale University Press.
- Oren, I. (2006). Can political science emulate the natural sciences? The problem of self-disconfirming analysis. *Polity*, 38, 72–100.
- Rabinow, P., & Sullivan, W. (Eds.). (1987a). *Interpretive social science: A second look*. Berkeley: University of California Press.
- Rabinow, P., & Sullivan, W. (1987b). The interpretive turn: A second look. In P. Rabinow & W. Sullivan (Eds.), *Interpretive social science: A second look* (pp. 1–32). Berkeley: University of California Press.
- Riley, D. (2007). The paradox of positivism. *Social Science History*, 31, 115–126.
- Schaar, J. H., & Wolin, S. S. (1963). Review essay: A critique [Review of the book *Essays on the Scientific Study of Politics*]. *American Political Science Review*, 57, 125–150.
- Shapiro, I. (2005). *The flight from reality in the human sciences*. Princeton, NJ: Princeton University Press.
- Skocpol, T. (2004). Voice and inequality: The transformation of American civic democracy. *Perspectives on Politics*, 2, 1–18.
- Steinmetz, G. (2005a). Introduction. In G. Steinmetz (Ed.), *The politics of method in the human sciences: Positivism and its epistemological others* (pp. 1–58). Durham, NC: Duke University Press.
- Steinmetz, G. (Ed.). (2005b). *The politics of method in the human sciences: Positivism and its epistemological others*. Durham, NC: Duke University Press.
- Storing, H. (Ed.). (1962). *Essays on the scientific study of politics*. New York: Holt, Rinehart & Winston.
- Strauss, L. (1962). An epilogue. In H. Storing (Ed.), *Essays on the scientific study of politics* (pp. 305–327). New York: Holt, Rinehart & Winston.
- Taylor, C. (1971). Interpretation and the sciences of man. *Review of Metaphysics*, 25, 3–51.
- Thiele, L. P. (1997). *Thinking politics: Perspectives in ancient, modern and postmodern political theory*. Chatham, NJ: Chatham House.
- Voegelin, E. (1987). *The new science of politics: An introduction*. Chicago: University of Chicago Press.
- Williams, R. (1985). *Keywords: A vocabulary of culture and society* (Rev. ed.). New York: Oxford University Press.
- Wolin, S. S. (1969). Political theory as a vocation. *American Political Science Review*, 63, 1062–1082.
- Yanow, D., & Schwartz, P. (Eds.). (2006). *Interpretation and method: Empirical research methods and the interpretive turn*. Armonk, NY: M. E. Sharpe.

56

CONSTRUCTIVISM

MARIYA Y. OMELICHEVA

University of Kansas

The last decades of the 20th century were marked by significant transformations on a global scale. The arrival of new forces created by discoveries in the realms of technology, transportation, and communications changed the patterns of social life and structures of international relations. The end of the cold war and ideological confrontation, decline in state sovereignty, and spread of globalization enlivened scholarly thinking about international relations and fostered academic debates about the nature of global politics and ways in which one can know and study it.

The arrival of constructivism in the late 1980s was precipitated by these earthshaking changes in international relations and lively discussions within the discipline. This novel heterodox approach imbibed the criticisms of the mainstream perspectives on international relations, particularly the theories of neorealism. The latter was faulted for its inability to account for changes in the global realm because of its neglect of the transformational power of knowledge and ideas. Instead of prioritizing the role of material factors in international relations, the constructivist perspective emphasized ideational forces. Instead of accepting relations and structures in global politics as the natural or given order of things, it maintained that a reality of international relations was contingent and dependent on people's thinking about it.

Beginning at the margins of the field, constructivist scholarship expanded rapidly throughout the 1990s. It developed its own program of empirical research focusing on identities, culture, institutions, knowledge, and norms. By shedding new light on the nature and impact of norms and ideas in international relations, interrogating identities and interests of states, and establishing ideational bases of the social order, constructivism has broadened theoretical confines of the discipline of international relations and contributed to reconceptualization of its key themes. Today, constructivist fortunes show no sign of waning, and the quality and depth of constructivist research has substantially improved.

The goal of this chapter is to take stock of constructivist work. It begins by explaining what constructivism is and laying out the main constructivist premises that hold this diverse perspective together and set it apart from other approaches within the discipline. Thinking about constructivism as a homogeneous approach obscures its rich philosophical roots that gave rise to numerous permutations within this approach. Therefore, the second section discusses various constructivisms and highlights divisions within the approach. Next, the chapter provides an overview of some of the seminal empirical works applying constructivist assumptions and methods, followed by a section discussing critical appraisals of the constructivist scholarship and directions for future research.

What Is Constructivism?

There is no unanimous agreement among scholars of international relations on what constructivism is. There is, however, broad consensus on what it is not. Constructivism is not a theory of international politics (Checkel, 1998; Finnemore, 1996; Wendt, 1999). It does not put forward general explanations for what individuals and states do, why societies differ, or how the world changes. Neither does it advance any claims about the content of international norms and institutions or the nature of participants of world politics. “Constructivism is empty as far as assumptions, propositions, or hypotheses about international relations are concerned” (Jørgensen, 2001, p. 41). What constructivism does offer is a set of ideas about the nature of reality and the ways in which it can be grasped, and these ideas can inform people’s understanding, interpretation, and theorization about world politics. In this way, constructivism can be thought of as an approach to studying social relations or a framework of propositions that lays the basis for social theories of international relations (Kratochwil, 2001).

Although in practice constructivist scholarship is very diverse and divided on a number of philosophical issues, most constructivists would agree that a defining aspect of this approach is the idea of the socially constructed nature of international politics. This idea encapsulates two inter-related processes: (1) The social environment makes individuals, states, and other actors of world politics into the kinds of beings and entities they are; and (2) conversely, individuals, states, and other actors of world politics make the world what it is through various forms of interaction with each other (Onuf, 1998).

First, for constructivists, the environment surrounding states and other actors of world politics is both social and material (Checkel, 1998; Jepperson, Wendt, & Katzenstein, 1996). The social world is composed of shared ideas and knowledge, whereas the material world manifests itself in the presence of nuclear weapons, the absence of world government, and other observable manifestations of international relations. However, the material aspects of world politics do not come classified. Material structures, beyond some biological characteristics, have certain meanings insofar as individuals and their collectivities create shared understandings of what those material structures signify, and individuals and groups attach this collective knowledge to physical reality (Adler, 2002). Things that individuals perceive as objective, such as money, human rights, or sovereignty, are made largely of ideas. They are the so-called social facts that depend on human agreement that they exist and make sense because people have imbued them with certain meanings (Ruggie, 1998).

Take, for example, human rights. They are the social constructions or inventions of the human mind that exist because of individuals’ beliefs in human rights and practices

reinforcing their existence (Schmitz & Sikkink, 2002). The norms of sovereignty or the institution of self-help have been perceived as natural and taken for granted. But they are nothing more than the artifacts of what people collectively believe in and practice (Wendt, 1992). These shared beliefs, understandings, knowledge, culture, and norms constitute an ideational context, which, according to constructivists, exerts a powerful impact on world politics because it defines the meanings of what individuals encounter and experience. Meaningful behavior in international relations is impossible without these shared understandings because people and states act toward others on the basis of meanings that they ascribe to them (Hopf, 1998; Wendt, 1992). States’ foreign policies toward other states, for example, will differ depending on whether their counterparts are perceived as enemies or friends. Nuclear weapons in and of themselves are less consequential for foreign policy choices than are our perceptions of states that possess them. The United States, for example, is not concerned with a sizable nuclear arsenal held by the British. However, North Korea’s aspiration to join a nuclear club is a major cause for alarm (Checkel, 1998; Wendt, 1992).

The effects of ideas penetrate deeper than states’ policies and behavior. Ideational context influences the basic character of states, the so-called state identity, the “relatively stable, role-specific understandings and expectations about self” (Wendt, 1992, p. 398). Actors’ identities tell them and others who they are and predispose them to embrace a particular set of interests and preferences over choices of action. An identity of great power furnishes a particular set of interests different from those implied by the identity of a European state. Because actors have multiple identities, constructivism does not accept the notion of fixed interests (Hopf, 1998).

Wendt (1999), for example, speculates that the international system of states can have at least three kinds of ideational contexts—Hobbesian, Lockean, or Kantian—distinguished on the basis of what kind of roles—enemy, rival, or friend—dominates the system. Each ideational context predisposes states to take a distinct position or orientation toward each other with respect to the use of violence. For Wendt, the contemporary system of states has a Lockean structure in which states assume role identities of rivals, recognize each other’s rights to life and liberty, and restrain their violence toward each other by observing the other’s right to exist.

Constructivists describe norms, beliefs, and knowledge that serve as the foundational blocks of the ideational context as *intersubjective*. The quality of intersubjectivity implies that meanings ascribed to social facts are not simply the aggregations of beliefs of individuals. Rather, they represent collective knowledge. This knowledge is created through dialogical relationships and interaction of actors (Fierke, 2001). The second premise of constructivism is

that the meanings in terms of which individuals' and states' actions are organized arise out of interaction (Wendt, 1992). By doing what they do and saying what they say, individuals create intersubjective meanings, thus making the world (Onuf, 1998). The repetition of these processes leads to the reproduction of intersubjective meanings that over time solidify and become objective social facts that are not easy to change or transform.

To recapitulate, in the constructivist worldview, international relations are inherently social. They are made of intersubjective understandings about the world as well as material objects. The intersubjective ideational structures define international actors, shape their identities, preferences, and interests and influence their behavior. The ideational structures themselves arise out of the actors' interaction. Individuals participate in the production, constitution, and fixing of the social reality through their actions, interactions, and discourse (Hopf, 1998; Wendt, 1999). The simultaneity of the processes of production of collective meanings and knowledge and the impact of the knowledge structures on actors' behavior and identities is known as the *mutual constitution* of agents and structures in the constructivist parlance.

The mutually constitutive relationship is one of the most challenging constructivist ideas to grasp. An easier way to comprehend it is through comparison with a more familiar causal relationship. Causal relationships are usually postulated in the form X causes Y or X leads to Y: Toxins cause cancer, fluctuations in crime rates cause changes in housing prices, and a plurality rule electoral system leads to the creation of a two-party system (Duverger, 1972). Causal relationships typically answer the why questions: Why did the housing prices drop? Why does one country have a two-party system, while another one has a multiparty system? By formulating answers to these types of questions in the form of a causal relationship between two phenomena, one typically assumes that the cause (X) and effect (Y) are independent of each other; the cause, temporarily, precedes the effect; and the latter would not have taken place without the former (Wendt, 1998).

Constructivists study social facts made of shared ideas and intersubjective understandings. Because they are interested in ferreting out what gives the cause (X) and effect (Y) certain meanings and how the relationship between X and Y came to be defined, the constructivists' goal is to account for the properties of social facts by referencing ideas and practices in virtue of which they exist (Wendt, 1998). This goal is accomplished by asking "How" and "What" questions (Wendt, 1998): How is it possible that chemical and nuclear weapons have become regarded as illegitimate instruments of warfare (Price & Tannewald, 1996)? What is Eurasian regionalism? What makes a region (Mansfield & Milner, 1999)? The answers to these questions are based on a different kind of logic that explains how various ideational factors—norms,

identities, culture, and knowledge—define what the social facts are, not what determines them. The factors in a constitutive relationship are not independent and separated in time: The factors constituting Europe or human rights neither exist apart from Europe or human rights nor precede them in time. Democratic culture, traditions, geography, and individual states do not cause Europe just as an international law of human rights does not cause human rights. They are constituted by intersubjective understandings, ideas, and beliefs about what Europe is and what human rights are.

Types of Constructivism

Thinking about constructivism as a homogeneous approach obscures the wide range of alternative conceptions of world politics and ways of studying it that exist under this rubric. There are numerous variants of constructivism—sociological, feminist, interpretive, emancipatory, and others. There are transnational constructivists who emphasize the influence of international norms, institutions, and other ideational structures (Boekle, Rittberger, & Wagner, 2001), and there are societal constructivists, also known as culturalists (Farrell, 2002), who stress the importance of domestic institutions, culture, and norms (Hopf, 2002; Katzenstein, 1996).

Although all variants of constructivism share their commitment toward denaturalization of the social world—that is, toward uncovering the socially constructed nature of institutions, objects, and practices that we perceive as objective (Hopf, 1998)—they disagree over the extent to which the empirically identifiable social relations can be discerned and studied (Barkin, 2003). One group, variously labeled as *thin*, *classical*, or *conventional* constructivists (Fierke & Jørgensen, 2001) to indicate that the group remains rooted in the classical or conventional tradition of viewing and studying international relations, maintains that an identifiable reality exists out there and can be examined and understood by applying appropriate methodologies for empirical investigation. Thin constructivists undertake to isolate sets of ideas, norms, and beliefs and specify a set of conditions under which one can expect to observe their impact on the behavior of states or other actors of international politics.

The other group, labeled as *thick*, *critical*, or *postmodernist* constructivists, contends that reality does not exist independently of individuals' or scholars' knowledge about it. It is apprehended in the form of multiple, intangible mental constructions derived from social experiences that may be specific or shared among many individuals and cultures. These constructions are neither false nor true; they are more or less informed and sophisticated (Guba & Lincoln, 1994; Schwandt, 1994). In thick constructivism, what exists (social reality out there) is entirely contingent on processes of social construction, in which an observer

inescapably takes part. Researchers contribute to the construction and reconstruction of reality through their scientific exploratory practices (Albert, 2001). They can never know for certain if what they observe really exists independently of their observation because the findings of social inquiry are literally created as the investigation proceeds. What can be known is inextricably intertwined with the interaction between an investigator and his or her object of study (Guba & Lincoln, 1994).

The bulk of constructivist scholarship can be characterized as conventional or thin constructivism (e.g., Checkel, 1997; Finnemore, 1996; Finnemore & Sikkink, 1998; Katzenstein, 1996; Wendt, 1999), albeit there are important differences and disagreements within this group. Some conventional constructivists grudgingly accept the dualism of a researcher and his or her object of inquiry and often use the language of causality and explanation in much the same way as positivists do. Others abandon dualism and think about human knowledge as being socially constructed and the relationship between an investigator and the objects of investigation as interactive with the values of an investigator inevitably influencing the inquiry. But instead of focusing on the matter of individual meaning-making activity of the individual mind as critical constructivists do, conventional constructivists concentrate on how people create intersubjective meanings and knowledge about the world in the process of social exchange (Schwandt, 1994).

Constructivist Empirical Research Program

Constructivist insights about the role of ideas in world politics provided an impetus for a variegated and rapidly expanding research program. Since constructivists are primarily interested in the social construction of international relations, their focus of inquiry has been on a range of social phenomena, such as norms, institutions, principled beliefs, culture, and knowledge.

Constructivists' studies of norms remain the staple of their scholarship, and a comprehensive list of the literature in this analytical realm would be impossible to compile. Researchers have documented the impact of separate norms, such as norms prohibiting colonization (Goertz & Diehl, 1994) and slavery (Ray, 1989), or sets of related norms, such as norms prohibiting certain types of conduct in the situations of war (Raymond, 1997) on different international outcomes, decolonization (Jackson, 1993), international support for the termination of slavery, and the emergence of weapons taboo (Price & Tannewald, 1996). Other scholars have demonstrated how the emergence of global standards of appropriate behavior, such as the norm of racial equality (Klotz, 1995), women's suffrage (Keck & Sikkink, 1998), and human rights (Forsythe, 1991) have led states to redefine their interests and change their behavior even in the absence of material incentives to do so.

The constructivist approach to norms differs from the perspectives on norms maintained by other theoretical approaches to international relations, particularly the treatment of norms within the realist and liberal schools of thought. Realists do not ascribe an independent causal role to a norm, while liberals treat norms instrumentally as tools for maximizing utility of the gain-seeking actors. In contrast, for constructivists, norms are independent forces, the effects of which reach much deeper than simply constraining states' behavior. Norms not only regulate the behavior of actors in international relations, but also create and define their very identities and interests (Checkel, 1997, 1998; Finnemore, 1996).

Before a norm reaches the status of an independent and constitutive force, it usually passes through several lengthy and uneven phases of its life cycle. Finnemore and Sikkink (1998) describe the pattern of the rise of a norm as a three-step process. The emergence of the norm is the initial stage. Norms are typically introduced and propelled by the norm entrepreneurs. The next stage is called the *norm cascade* when state actors begin to adopt the norm. This process culminates in a tipping point, at which a sufficient number of the relevant actors accept the norm. The last stage involves norm internalization. Risse, Ropp, and Sikkink (1999) suggest an alternative spiral model that outlines five steps in norm adoption. During the early stages of the norm adoption cycle, a government does not acknowledge the existence of the norm and denies violations but may be pressed into tactical concessions by domestic and international social actors demanding compliance with the norm. In the last two stages of the cycle, the norm reaches a prescriptive status and, ultimately, becomes internalized by the state actors leading to their rule-consistent behavior.

Different stages of a norm's life cycle and steps in the adoption of the norm may involve different actors. Stages of a norm's life cycle are also characterized by different social processes, logics of action, and causal mechanisms connecting driving forces for the norm's emergence and adoption with certain outcomes. For example, constructivist studies illuminated the role of norm entrepreneurs—the networks of activists, knowledge-based experts, and “epistemic communities” (Finnemore & Sikkink, 1998)—at the initial stages of norm emergence and progression from the local to global level. The key processes at the early phase of the life cycle of norms are framing that aims to make the norm appealing and comprehensible to the public and engaging in persuasion to convince the leaders of states and international organizations to embrace new norms. When the norm reaches the global arena, it is the international organizations and like-minded states that serve as an organizational platform for advocating the norm and teaching new normative views to other states (Finnemore, 1996). They can exert international pressure or use legitimation as mechanisms for socializing the states into becoming the norm followers (Finnemore & Sikkink, 1998).

The life cycle of a norm, or phases of norm adoption, is not a linear process. Internalization of the norm is never a guaranteed outcome. Constructivist scholarship offers an impressive catalogue of what are called the *scope conditions* under which norms are likely to fail or succeed (Zürn & Checkel, 2005). Some of these conditions encompass the properties of the norms themselves, such as norms' specificity, commonality, and durability (Boekle et al., 2001; Raymond, 1997), while others apply to the properties of actors and institutions that trigger socialization. Another group of conditions includes the properties of domestic political systems—the nature of political regimes, the strength of civil society, and the lines of political contestation—that can either facilitate or obstruct the implementation of norms. Finally, the content of issues (for instance, human rights or democratic governance) and the nature of interaction between socializing and socialized actors (for instance, intensity of contact and discourse) also condition the impact of norms (Zürn & Checkel, 2005).

Constructivist scholarship is not limited to norms. The burgeoning analyses of individual and collective identity, how it is created and sustained, and how it generates and shapes interests and policies of international actors constitute a big chunk of constructivist research (Cronin, 1999; Jepperson et al., 1996). Constructivists have broken new ground in the studies of institutions—that is, relatively stable collections of rules and practices prescribing and proscribing certain kinds of behavior for a group of actors (March & Olsen, 1998)—by showing how institutions not only help to coordinate, pattern, and direct behavior of states, but also partake in the creation of new collective identities, definition of shared interests, and promotion of new practices. Finally, constructivist analyses of nonstate actors and issues of international governance have made a substantive impact on the international relations discipline (Adler, 2002).

Criticisms of Constructivism and Future Directions of Constructivist Scholarship

Constructivism has been subjected to scrupulous internal and external evaluation. Scholars within and outside constructivism have found important limitations and shortcomings of multiple substantive theories and empirical studies informed by this approach. The major criticisms of constructivism originate from those theoretical perspectives that fall under the rubric of positivism. The latter prioritizes causal laws and generalizations describing and explaining the reality that is assumed to be independent of people's thinking about it, even if this thinking is never complete and perfect (Fierke, 2001). Positivist tenets underlie the mainstream perspectives on international relations, such as realism and liberalism.

The critics of constructivism contend that its usefulness as a guide for studying international relations is limited. Theories informed by constructivist assumptions are not

parsimonious or elegant, their causality is indeterminate, and relationships are not clearly specified. Constructivists devise cumbersome models including different actors and describe complex mechanisms of influence and scope conditions that are difficult to apply beyond the situations and processes under their investigation.

A constructivist idea of the mutually constitutive relationship between actors and structures has become a target of many attacks. Constructivists have been faulted for their inability to disentangle the mutually constitutive relationships and establish their temporal sequence: What comes first, a norm that affects the identity of actors or actors' identities that influence the nature of norms? The simultaneity of interaction makes it very difficult to capture the self-reinforcing nature of norms, institutions, or cultures and the ways in which states, individuals, and other social agents create and change the social order of things.

Grounding their explanations in unobservable (inter-subjective) ideational structures, constructivists have to tackle two formidable methodological challenges. First, they need to demonstrate the existence of norms, and second, they need to prove their impact on the behavior of states (Farrell, 2002). To show the existence of shared beliefs, constructivists rely on the artifacts of actors' interactions, such as public statements, decisions of authoritative bodies, or official memoirs. The residues of the culture and norms have also been found in international and domestic legislation. To tease out the meanings that actors ascribe to social facts and situations, constructivists have employed interpretive methods and a narrative mode of explanation (Klotz & Lynch, 2007) that have been regarded as less methodologically robust tools of research.

Another complaint about the constructivist agenda is that it has tended to be liberal idealist, concentrated on Western liberal norms of democracy, human rights, or multilateralism (Keck & Sikkink, 1998; Risse et al., 1999). Although constructivists have begun examining the so-called bad norms and pathological identities (Farrell, 2002; Rae, 2002), their research has overwhelmingly focused on so-called good norms. One of the implications of this selection bias is the erroneous representation of the West and Western organizations as promoters of good liberal norms that stimulate progress in international relations.

Excessive emphasis on the ability of the good norms and other ideational factors to change the world and insufficient attention to material coercion and political contestation in world politics has created an image of constructivism as an approach dismissive of the role of power in the creation and dissemination of norms and ideas (Barkin, 2003). By ignoring or downplaying the advantages that material resources and power give to some social actors of international relations, constructivists overlook significant interrelated effects of social and material inequalities on the nature, patterns of diffusion, and ultimate success of international practices and norms.

The ambivalence toward or neglect of the role of power structures in international relations by some constructivists

can, to a certain extent, be attributed to insufficient attention to domestic politics and the lack of a theory of agency in constructivist research. Constructivism emerged on the wave of the growing dissatisfaction with the neorealist individualistic and systemic orientation. Yet it has been conspicuously inattentive to the state-level accounts of world politics. Certainly, there are constructivists who attend more closely to domestic power constellations and culture as mediating factors in the adoption of norms or domestic sources of foreign policies and international relations (see, for example, Checkel, 1997, 1998; Hopf, 2002). However, the bulk of constructivist scholarship has remained at the international level of analysis continuing to treat states as unitary actors.

Most of the failings identified by constructivism's critics are not terminal. They can be cured in future research that should respond to the needs of theory building and greater attentiveness to the role of power in the social world. As stated previously, constructivism is not a theory of international relations. For it to serve as a valuable guide into the exploration of the social world, its abstract philosophical categories and insights about the nature of social relations need to be translated into the middle-range theory with a more limited scope and aiming at explaining a set of specific social phenomena. Constructivist scholarship has seen laudable efforts to formulate and test middle-level theories specifying the actors and mechanisms of social influence and articulating conditions under which social influence occurs. There is still an unfortunate deficit of constructivist theory building in international relations, and there is a lack of conversation among constructivists of different genres (Checkel, 1998). Future studies inspired by constructivist propositions need to elaborate the causal pathways and transmission mechanisms that link norms, actors, and their policy choices in various social situations. There is also room for specifying the meaning of concepts and relationships and detailing conditions under which different mechanisms of normative influence can be observed.

Opening up the black box of domestic politics for theoretical and empirical exploration and attending more closely to the structures of power in domestic politics and international relations will facilitate constructivists' efforts at theory building and enhance their explanations and reconceptualizations of practices and structures in the international realm. Having demonstrated the importance of ideational forces in world politics, constructivists need to contemplate how and why certain norms and beliefs get successfully diffused, promoted, and adopted by international actors but others do not. What is the relationship between social and material power, how is it wielded, and to what end? Future constructivist studies should consider not only the impact of ideational factors on the structures and exercise of power in the international realm, but also the ways in which power and political contestation in domestic politics and international relations influence and condition the impact of norms (Barkin, 2003). The

accomplishment of these goals will require a synthesis of constructivist and other theoretical approaches and bridge building between constructivism and rationalism. Efforts at integrating constructivism with other theoretical perspectives are well under way (Barkin, 2003; Checkel, 1997; Risse et al., 1999), but there are obstacles toward the bridge building and reconciliation.

Furthermore, future constructivist studies should expand their research agenda to include norms, institutions, and identities that are not accepted as good and critically assess ethical implications of the diffusion of good norms. Constructivist scholars need to refine their research designs and hone their methods of empirical investigation. Much of the empirical constructivist work has focused on examining single countries or issues. Cross-national or longitudinal designs as well as considerations and tests of alternative explanations would help to reduce the problem of overdetermination that is evident in much of the constructivist research, where ideational factors are invoked as one of the explanatory factors, yet little consideration is given to other variables and how much of the outcome they can account for (Checkel, 1998).

Conclusion

Constructivism in international relations is a fairly new approach that focuses on the social construction of world politics. It emphasizes ideational factors, such as ideas, beliefs, and knowledge, and their constitutive and regulative effects on the social reality and agents that create, reproduce, and reify it. Intersubjective ideational contexts influence actors' behavior and identities by embedding material objects, including other actors, with which they interact with certain meanings, and those shared meanings become the source of the agents' reasons, interests, and practices. Social actors themselves create intersubjective meanings through their discourse and interactions. For constructivists, a reality is always the product of human activity; therefore, it is never objective or given but is always historically bound and contingent. Constructivist scholarship is extremely variegated and divided along philosophical, theoretical, and methodological issues. However, all constructivists share their commitment toward denaturalization of the world—that is, discerning how material objects, practices, and institutions that individuals treat as given and natural are the products of social construction (Hopf, 1998).

One of the important strengths of the constructivist approach is its capacity to account for what the mainstream theoretical perspectives cannot—namely, change in the structures and agents of international politics, including visible shifts in the goals, behaviors, and strategies of states (Locher & Prügl, 2001). For constructivists, preferences and interests are the products of human activity; therefore, they can change with instantiation of new social practices, although this process can be incremental and

slow (Wendt, 1999). For realists and liberals, qualitative changes in interests and goals are difficult to explain because they are postulated as exogenous to the actors, and therefore, not variable.

Constructivist scholarship is not devoid of limitations and shortcomings. Future constructivist studies need to pay closer attention to the development of theory, the mediating role of power in the emergence and diffusion of ideas and norms, and the design of the empirical research. Despite the criticisms, constructivism's contributions to international relations cannot be underestimated. The constructivist approach has significantly broadened theoretical and empirical contours of the discipline. It has improved understanding of some of the conceptual foundations of international relations theory and suggested novel ways of thinking about key themes and concepts in international relations, such as anarchy, balance of power, and the security dilemma, to name a few (Hopf, 1998). By attending to the issues of identity and construction of interests bracketed by mainstream theoretical perspectives, constructivist studies have put forth alternative interpretations of international phenomena and offered new solutions for a number of puzzles of international relations (Checkel, 1998). Constructivism's empirical research on principled beliefs, culture, knowledge, and norms has filled in a clear lacuna in the contemporary international relations literature.

References and Further Readings

- Adler, E. (2002). Constructivism and international relations. In W. Carlsnaes, T. Risse, & B. Simmons (Eds.), *Handbook of international relations* (pp. 95–119). Thousand Oaks, CA: Sage.
- Albert, M. (2001). What systems theory can tell us about constructivism. In K. Fierke & K. Jørgensen (Eds.), *Constructing international relations: The next generation* (pp. 93–114). Armonk, NY: M. E. Sharpe.
- Barkin, S. (2003). Realist constructivism. *International Studies Review*, 5, 325–342.
- Boekle, H., Rittberger, V., & Wagner, W. (2001). Constructivist foreign policy theory. In V. Rittberger (Ed.), *German foreign policy since unification* (pp. 105–137). Manchester, UK: Manchester University Press.
- Checkel, J. (1997). International norms and domestic politics: Bridging the rationalist constructivist divide. *European Journal of International Relations*, 3, 473–495.
- Checkel, J. (1998). The constructivist turn in international relations theory. *World Politics*, 50, 324–348.
- Cronin, B. (1999). Community under anarchy: Transnational identity and the evolution of cooperation. New York: Columbia University Press.
- Duverger, M. (1972). Party politics and pressure groups: A comparative introduction. New York: Thomas Y. Crowell.
- Farrell, T. (2002). Constructivist security studies: Portrait of a research program. *International Studies Review*, 4, 49–72.
- Fierke, K. (2001). Critical methodology and constructivism. In K. Fierke & K. Jørgensen (Eds.), *Constructing international relations: The next generation* (pp. 115–136). Armonk, NY: M. E. Sharpe.
- Fierke, K., & Jørgensen, K. (Eds.). (2001). *Constructing international relations: The next generation*. Armonk, NY: M. E. Sharpe.
- Finnemore, M. (1996). *National interests in international society*. Ithaca, NY: Cornell University Press.
- Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization*, 52, 887–917.
- Forsythe, D. (1991). *The internationalization of human rights*. Lexington, MA: Lexington Books.
- Goertz, G., & Diehl, P. (1994). International norms and power politics. In F. Wayman & P. Diehl (Eds.), *Reconstructing realpolitik* (pp. 101–125). Ann Arbor: University of Michigan Press.
- Guba, E., & Lincoln, Y. (1994). Competing paradigms in qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–118). Thousand Oaks, CA: Sage.
- Hopf, T. (1998). The promise of constructivism in international relations theory. *International Security*, 23, 171–200.
- Hopf, T. (2002). *Social construction of international politics: Identities and foreign policies, Moscow, 1995 and 1999*. Ithaca, NY: Cornell University Press.
- Jackson, R. (1993). The weight of ideas in decolonization: Normative change in international relations. In J. Goldstein & R. Keohane (Eds.), *Ideas and foreign policy* (pp. 11–38). Ithaca, NY: Cornell University Press.
- Jepperson, R., Wendt, A., & Katzenstein, P. (1996). Norms, identity, and culture in national security. In P. Katzenstein (Ed.), *The culture of national security: Norms and identity in world politics* (pp. 33–78). New York: Columbia University Press.
- Jørgensen, K. (2001). Four levels and a discipline. In K. Fierke & K. Jørgensen (Eds.), *Constructing international relations: The next generation* (pp. 36–53). Armonk, NY: M. E. Sharpe.
- Katzenstein, P. (1996). *Cultural norms and national security: Police and military in postwar Japan*. Ithaca, NY: Cornell University Press.
- Keck, M., & Sikkink, K. (1998). *Activists beyond borders*. Ithaca, NY: Cornell University Press.
- Klotz, A. (1995). *Norms in international relations: The struggle against apartheid*. Ithaca, NY: Cornell University Press.
- Klotz, A., & Lynch, C. (2007). *Strategies for research in constructivist international relations*. Armonk, NY: M. E. Sharpe.
- Kratochwil, F. (2001). Constructivism as an approach to interdisciplinary study. In K. Fierke & K. Jørgensen (Eds.), *Constructing international relations: The next generation* (pp. 13–35). Armonk, NY: M. E. Sharpe.
- Locher, B., & Prügl, E. (2001). Feminism and constructivism: Worlds apart or sharing the middle ground? *International Studies Quarterly*, 45, 111–129.
- Mansfield, E., & Milner, H. (1999). The new wave of regionalism. *International Organization*, 53, 589–627.
- March, J., & Olsen, J. (1998). The institutional dynamics of international political orders. *International Organization*, 52, 943–969.
- Onuf, N. (1998). Constructivism: A user's manual. In V. Kubalkova, N. Onuf, & P. Kowert (Eds.), *International relations in a constructed world* (pp. 58–78). Armonk, NY: M. E. Sharpe.

- Price, R., & Tannenwald, T. (1996). Norms and deterrence: The nuclear and chemical weapons taboos. In P. Katzenstein (Ed.), *The culture of national security: Norms and identity in world politics* (pp. 114-152). New York: Columbia University Press.
- Rae, H. (2002). *State identities and the homogenization of people*. New York: Cambridge University Press.
- Ray, J. (1989). The abolition of slavery and the end of international war. *International Organization*, 43, 405-439.
- Raymond, G. (1997). Problems and prospects in the study of international norms. *Mershon International Studies Review*, 41(Suppl. 2), 205-245.
- Risse, T., Ropp, S., & Sikkink, K. (Eds.). (1999). *The power of human rights: International norms and domestic change*. New York: Cambridge University Press.
- Ruggie, J. (1998). What makes the world hang together? Neo utilitarianism and the social constructivist challenge. *International Organization*, 52, 855-885.
- Schmitz, H., & Sikkink, K. (2002). International human rights. In W. Carlsnaes, T. Risse, & B. Simmons (Eds.), *Handbook of international relations* (pp. 517-537). Thousand Oaks, CA: Sage.
- Schwandt, T. (1994). Constructivist, interpretivist approaches to human inquiry. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118-138). Thousand Oaks, CA: Sage.
- Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 46, 391-425.
- Wendt, A. (1998). On constitution and causation in international relations. *Review of International Studies*, 24, 101-118.
- Wendt, A. (1999). *Social theory of international politics*. Cambridge, UK: Cambridge University Press.
- Zürn, M., & Checkel, J. (2005). Getting socialized to build bridges: Constructivism and rationalism, Europe and the nation state. *International Organization*, 59, 1045-1079.

57

REGRESSION ANALYSIS

MARK ZACHARY TAYLOR

Georgia Institute of Technology

This chapter provides a brief introduction to regression analysis. Regressions are a form of statistical analysis frequently used to test causal hypotheses in social science research. A simple way of thinking about regressions is to try, given a scatterplot of data, to fit the best line to run through, and thereby describe, those data. Regressions are useful because that line can tell researchers a lot more about whether the data support a hypothesis than just the scatterplot alone.

Research papers and articles using regressions usually have a typical format. First, the author poses a research question or causal hypothesis. He or she then reviews recent debate and research regarding this question. Next, the author suggests a statistical regression model and data with which to help answer the research question (i.e., test the causal hypothesis) and thereby advance the scientific debate. The big scientific payoff comes in discussing the results of the statistical analysis, which usually includes a few tables of regression results. Finally, conclusions are drawn, and perhaps implications are suggested for policy or future research.

This chapter provides an introductory explanation of what bivariate and multivariate regressions are and how they work. The basic method of calculating regressions is called *ordinary least squares* (OLS), which is the focus of this chapter. OLS is the four-door sedan of statistics. It is the technique that most researchers prefer to

use. Certainly there are more complex and esoteric techniques, such as probit, logit, scobit, distributed lag models, and panel regression.¹ However, scientists use these techniques only in those special cases where OLS does not work so well. Therefore, this chapter also discusses briefly the conditions under which OLS can fail and offers some fixes within OLS that researchers use to address these failures.

Drawing a Regression Line: The Basics of Bivariate Regression

Given knowledge of introductory statistics (i.e., descriptive statistics, probability, and statistical inference), a student's next step is typically to take intermediate statistics, which for political scientists is always regressions. Why? Because testing theories of causality is the main goal of political scientists, who hypothesize that X causes Y; then, to test this hypothesis, they gather data and use regression analysis to see whether those data show any evidence of a causal relationship between X and Y.

Couldn't we just use introductory statistics to ask whether X and Y correlate or covary? Sure we could, but it would not tell us a whole lot. Let's see why, with an example of a bivariate regression, which is a regression

that has only two variables: one independent and one dependent.

Say the city newspaper reports that a strange flu has broken out in town, and it prints a chart of the number of sick people by zip code. You talk to a few flu sufferers, ask what they did during the days leading up to the flu, and find that they have one thing in common: They each went to meet with the city's mayor. So you ask yourself why that should cause the flu. Then you realize that the mayor is one of those old-style, glad-handing politicians whose style is to handshake and hug everyone he or she meets. So you hypothesize that the flu is caused by a virus that the mayor has and he or she is passing it on by this contact. But if you called the mayor's office or state health officials with this claim, the officials would think you were crazy. To better convince them, you need to provide some evidence to support your hypothesis. How can you do this?

You certainly cannot contact everyone in the city, asking if they visited the mayor recently and now have the flu. However, you could consult the mayor's official visitors' sign-in book, which records each visitor's name and address. Therefore, for each zip code, you could plot the number of visitors to the mayor on one axis and, using the newspaper's

chart, the incidence of flu on the other. It's simple; just draw a scatterplot.

Figure 57.1 shows a few possible scatterplots that might result from these data. Each scatterplot represents just one of an infinite set of possible results. If your data produced any of the first three scatterplots, then you could be pretty sure there is a relationship between visits to the mayor's office and incidence of the flu. But in the fourth scatterplot, you are not so sure. Therefore, the first thing you would like is a technique by which you could be more confident of whether you are observing a relationship in scatterplot 4. Descriptive statistics, like covariation and correlation, will not help much with this.

Also, look at the first three scatterplots. Each graph clearly shows some sort of a relationship between visits to the mayor and incidence of flu, but they are distinctly different relationships. It would be nice if you could say something about how they are different. Again, descriptive statistics are not much help here.

Let's start by figuring out how the relationships are different in scatterplots 1, 2, and 3. One technique is to simply draw a line through the data. How does drawing a line help? Remember from elementary algebra that given a

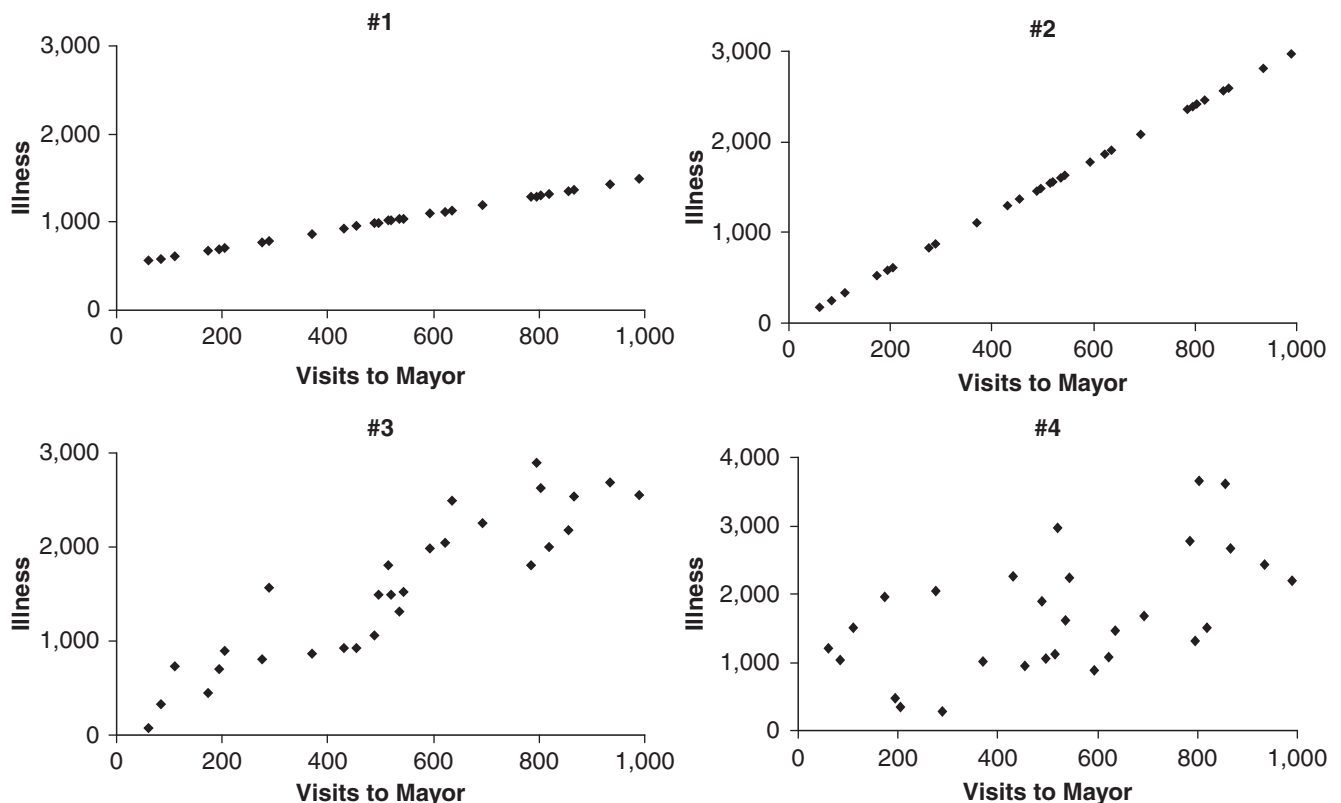


Figure 57.1 Illness Versus Visits to the Mayor

bunch of Y 's and X 's defined as points on a line, the equation for that line is

$$Y = (\text{slope} * X) + \text{intercept.} \quad (1)$$

First, the slope of the line will tell how much of an increase in Y is related to an increase in one unit of X . Second, the intercept of the line with the y -axis will tell how much Y there will be when there is no X at all (when $X = 0$).

So let's try to draw some lines through our mayor visits versus illness data. The first two are easy since the data points are perfectly lined up. In scatterplot 1, you would use simple algebra on the raw data (not shown) to figure out that the slope is 1, and the intercept is 500. Therefore, the result is Figure 57.2.

In scatterplot 2, you again use the raw data to calculate that the slope is 3, and the intercept is 0. Therefore, the result is Figure 57.3.

Now that we know these simple equations, we have two drastically different interpretations of the data. In scatterplot 1, the slope is 1. This means that for every one visit to the mayor, one person gets sick. Also, the intercept is 500; this implies that if no one visited the mayor, then 500 people would be sick. So the line in scatterplot 1 does not sound like a flu virus. That is, if someone visits the mayor and catches the flu, then he or she should bring it home to infect his or her friends, family, and then coworkers. So every one visit to the mayor should probably result in multiple people getting sick. In other words, the slope should be higher, which is exactly what one sees in scatterplot 2.

Furthermore, if the mayor is the source of the flu, then the following should apply: If no one visits the mayor, then few people should get sick. That is, the intercept term should be close to zero. Again, this looks more like what

we see in scatterplot 2. But in scatterplot 1, the line predicts 500 sick people even when no one visits the mayor.

Therefore, the equation for the line drawn in scatterplot 2 looks like evidence of a virus: If no one visits the mayor, no one gets sick, and for every person who visits the mayor, several people get sick. Meanwhile, the equation for the line drawn in scatterplot 1 does show that visits to the mayor's office correlate with illness, but this relationship does not look like a virus. So if your research produced scatterplot 1, then you might have to reject the virus hypothesis and formulate another one. For example, maybe it's the bad coffee they serve, moldy walls in city hall, or some sort of noxious gas at that particular subway stop. In this case, you would then want to gather new data (e.g., on the incidence of coffee drinking, mold allergies, etc.) to better test these new hypotheses.

This is exactly how political scientists use regressions. We first ask, "What causes Y to vary?" Then we formulate a hypothesis in which we theorize an X that causes Y to vary.² Then we gather data on X and Y and use regression analysis to draw a line through the data. Finally, we ask whether the slope, the intercept, and perhaps the shape of the line supports our hypotheses about what's going on between X and Y .

In a typical undergraduate regressions course, students might practice calculating some of the underlying mathematics by hand. But in practice, statistical software packages such as STATA, SPSS, SAS, R, Eviews, and dozens of others do the mathematical work for you. Data can be directly entered into these programs or imported from a typical spreadsheet program.³ To perform a regression, one need only type a few simple commands and then examine the computer readout. Table 57.1 shows part of a readout typical of that provided by many statistical software packages. This readout may look ugly, but it is just computereese

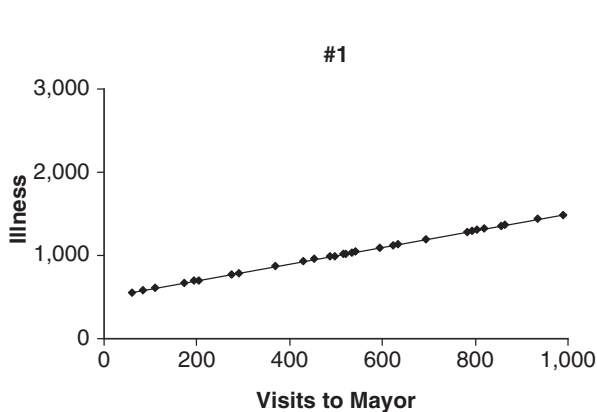


Figure 57.2 Scatterplot 1

NOTE: Sick people = 1 * [mayor visits] + 500

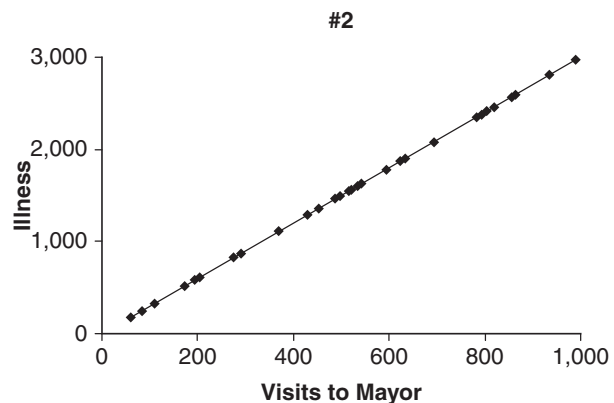


Figure 57.3 Scatterplot 2

NOTE: Sick people = 2 * [mayor visits] + 0

Command: regress illness visits						
Number of obs = 29						
R squared = 0.84						
illness	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]	
visits	2.68	0.23	11.76	0.00	2.21	3.15
cons	152.63	133.51	1.14	0.26	-121.25	426.51

Table 57.1 Sample Readout of Regression in Scatterplot 3

for scatterplot 3. This chapter now discusses how to make sense out of it. (see Figure 57.4)

First, where in the computer readout is the line's equation? You can find it in the lower left quadrant of the readout. The *Y*, or dependent variable (*illness*), is listed at the top of the column, with the independent variable (*visits*) and intercept term (*cons* for constant) below it. The values of the slope and intercept term are listed in the column labeled *Coef.* (for coefficient). Therefore, the readout tells that the line that best fits the data is described by $\text{illness} = 2.68 (\text{visits}) + 152$. We get two pieces of information from this equation:

1. For every 1 additional visitor to the mayor, an average of 2.68 people get sick.
2. When nobody visits the mayor, an average of 152 people get sick.

Note that we do not interpret these numbers as follows: One visit causes 2.68 cases of illness. Why? Remember that regressions can show correlation but not causality. Social scientists use statistics to argue for causality by asking whether the correlations they show match those predicted

by our hypotheses. Thus, statistics rarely prove a hypothesis, but they can provide evidence for or against it. Hence, the researcher's most important task is to identify the regression equation and data that will best test his or her hypothesis and thereby best convince an audience of fellow scientists that it is true.

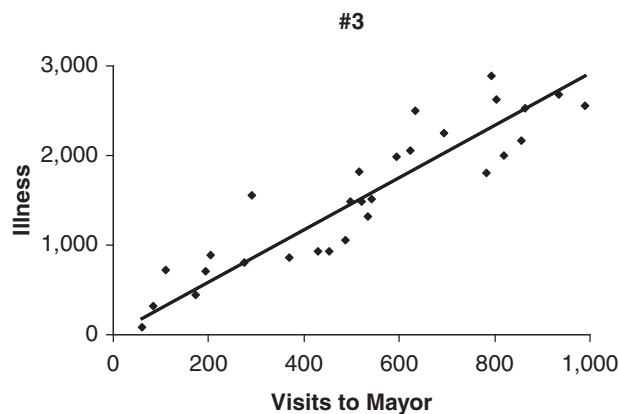
But what about all of the other information presented in the computer readout?

The Coefficient of Determination (R^2)

R^2 (or *R squared*) is known as the *coefficient of determination* or the *goodness of fit test*. Its value can range from 0 to 1; it tells you how linear the data are. What does this mean exactly? Let's look again at our four preceding scatterplots, this time with the OLS regression lines drawn through them (Figure 57.5). It may not look like it, but the relationships estimated in scatterplots 2, 3, and 4 are all based on the same underlying equation: $y = 3x$; hence, the true regression line's slope is 3, and the true intercept is 0. But there is something clearly different going on in each of them. Let's investigate.

In scatterplot 2, the line goes through each of the data points. That is, the regression line completely fits each and every data point; it therefore explains all the data. This means that all the variation in *Y* is explained by *X*; all the variation in illness is explained by visits to the mayor.

But in scatterplots 3 and 4, the data points wander or err on either sign of the line. That is, the regression line explains some of the data (the variation in *X* explains some of the variation in *Y*), but there is some left over, unexplained, or residual variation. More precisely, for any data point not on the regression line, part of its height is explained by the line, and part is not. The percentage of (height)² that is explained by the regression model tells us the percentage of *Y* explained by *X*. That percentage is known as R^2 and is reported as a number from 0.00 to 1.00. In scatterplots 1 and 2, the $R^2 = 1.00$ because *X* explains all

**Figure 57.4** Scatterplot 3: With Regression Line
($y = 2.68x + 152$)

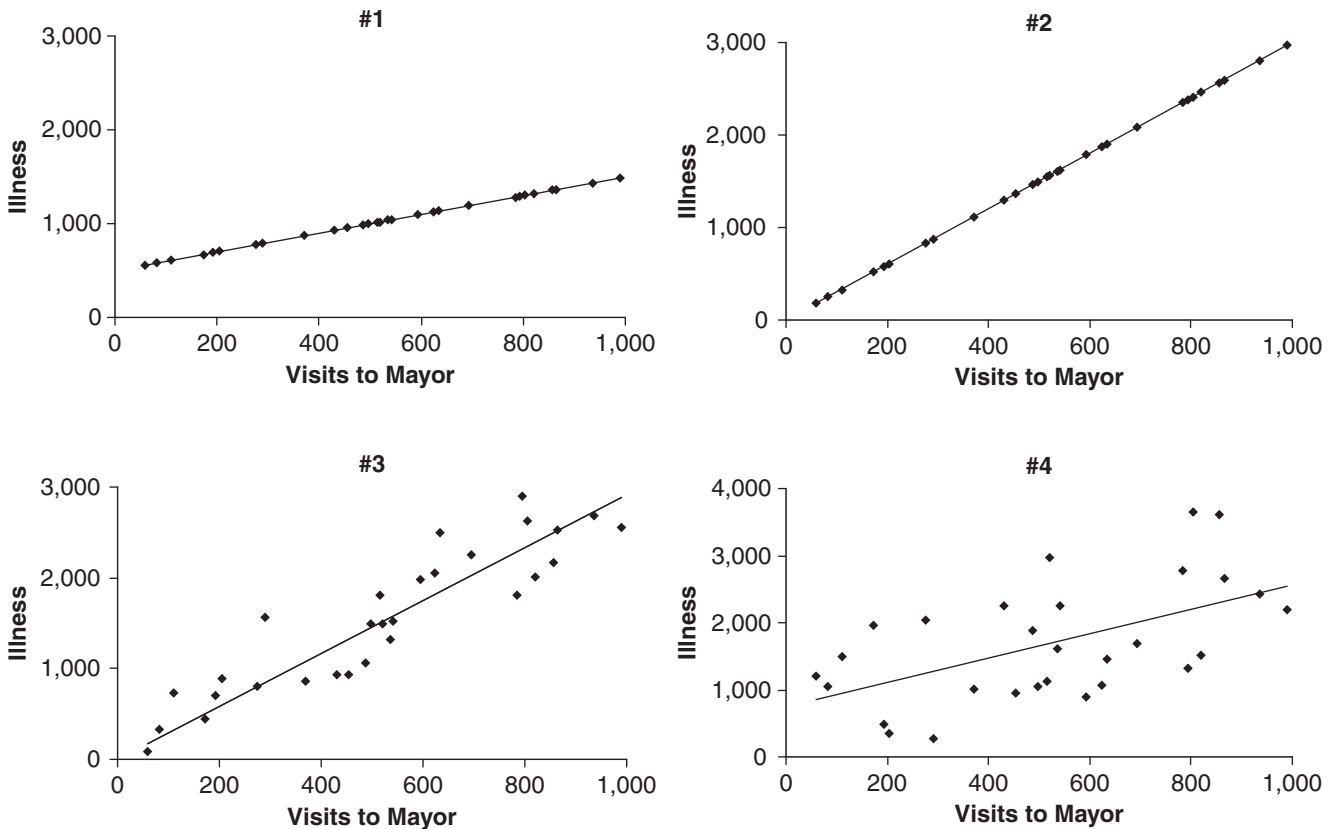


Figure 57.5 Illness Versus Visits to the Mayor (With Regression Lines)

of the variation in Y ; there is no unexplained portion of the data. In scatterplot 3, the $R^2 = 0.84$, which means that 84% of the variation in illness is explained by visits to the mayor. In scatterplot 4, the $R^2 = 0.31$, which means that, for these data, 31% of the variation in illness is explained by visits to the mayor.

What's behind the unexplained part of the data? It might simply be randomness. For example, people in one zip code may have been accidentally rubbing virus into their eyes more because of a random spike in air pollution this month, or the weather in another zip code was randomly warmer, making flu resistance slightly higher there.

However, there could also be a systematic cause for some of the unexplained data. Let's say that Norwegians are genetically more susceptible to this mystery flu, and the neighborhoods represented by data points above the regression line have more Norwegians than those below. The zip codes with more Norwegians would then have higher incidence of flu than one could explain using just visits to the mayor. And if we included another X in the regression, in which X = the percentage of the population that's Norwegian, we would get a higher R^2 because that part of the variation in Y would be explained. We will talk

a little more about adding more X 's and what that means in a subsequent section, but first let's finish up with R^2 .

A small warning about R^2 is appropriate here. Decades ago, eager young researchers used to jump on this R^2 measure and conclude something like the following:

Aha! Regressions 1 and 2 are better than 3 because they have higher R^2 ; likewise, Regression 3 is better than 4 because it has a higher R^2 . It explains more of the variation in the dependent variable.

But that is not entirely accurate. Remember that the true equation that generated the data ($y = 3x$) is the same for scatterplots 2, 3, and 4. The only difference is that the latter scatterplots have higher levels of randomness added to the data. So the exact equation for these scatterplots is this: $y = 3x + \epsilon$, where ϵ is some small random number. But the fundamental relationship between X and Y is the same across the scatterplots; therefore, one regression line is not somehow better than the other. Yes, in Regressions 3 and 4, there is a large amount of unexplained variation that is not present in 2. But that does not necessarily make the former bad regression lines. After all, Regressions 3 and 4 still show a relationship between

X and Y that indicates that a virus is present: high slope, low intercept.

So what does R^2 really tell you? It does not tell you whether you have a relationship between X and Y or whether one regression estimate is somehow better than another. What it tells you is how closely grouped around the regression line the data are. It tells you how linear the relationship is. For example, if we had a U-shaped scatterplot, then OLS would produce a low R^2 , since it would not be able to fit a straight line to the U of data. But that low R^2 does not mean that there is no relationship at all, just that there is no linear relationship. So although you generally cannot use R^2 to say that one regression is better than another, you might use R^2 to judge whether the data are linear, whether there are some missing independent variables causing a lot of systematic error (wandering), or whether there is a large component of randomness affecting your dependent variable.

Standard Errors of Coefficients

The most important information to come out of any regression is often the slope of the regression line (aka the coefficient) and perhaps its intercept. These are the two pieces of analysis that really tell us something useful about the relationship between X and Y . For example, remember our four preceding scatterplots. In them, the slopes, and to a lesser extent the intercepts, provided evidence as to whether we were dealing with a contagious virus.

But the numbers that OLS produces for the slope and intercept are just statistical estimates. Again, recall that the city has a large population, split up into dozens and dozens of zip codes. In our example, we merely took a sample of them and used this small sample to estimate the unknown parameters of the entire population. But how confident can we be in our estimates? In our example, how confident can we be that the true relationship, the relationship for everyone in the city, is actually $\text{illness} = 3 (\text{visits}) + 0$?

One important indicator of how far off the estimate might be is the standard error of the slope coefficient. The standard error is simply the standard deviation of how much the data wander around the regression line. It can therefore give us an indication of how much that point estimate is likely to vary from the true value.

You can use the standard error of the slope coefficient in the same way that you used standard deviations to judge sample averages in introductory statistics. In introductory statistics you should have learned that if a sample is selected at random, then as you increase the sample's size, the mean and variance of that sample will look more and more like the mean and variance of the population from which that sample was drawn. Furthermore, we know that for data that are normally distributed, 68% of the data will lie within 1 standard deviation of the mean, 95% of the

data will lie within 2 standard deviations, and so on. And thanks to the central limit theorem, we know that many important statistics, like means and variances, are normally distributed.

The slope and intercept terms that we estimate in regression analysis are no different. We can measure how much our data wander, and use this information to get a sense of how accurate our estimates of the slope and intercept are. Because of the central limit theorem, we can say that, yes, the slope and intercept we estimate from our sample may be a little higher or lower than that of the population from which our sample came. But 95% of the time, the estimates based on our sample data will wander within plus or minus 2 standard deviations of the population's values. Thus, we can be 95% confident that the population's slope and intercept lie within 2 standard errors (technically 1.96) of the coefficients we estimated from the sample.

For example, if our slope estimate is 3 and the standard error (the standard deviation of the wandering of the data) is 0.5, then we can be 95% confident that the true slope is 3, plus or minus 2 standard errors. That is, the sample came from a population with a regression slope between 2 and 4. This means that we are 95% confident that there is a positive (i.e., nonzero) relationship between visits and illness. We therefore say that the coefficient on visits to the mayor is statistically significant, given a 95% confidence level. In this particular case we can go further and say that the slope is greater than 1, that this relationship is greater than 1 to 1 (between 2 to 1 and 4 to 1). Hence, if a virus hypothesis predicts a relationship greater than 1 to 1, then our data support this hypothesis.

However, if our slope estimate is 3 and its standard error is 2, then we can be 95% confident that the population's slope is between -1 and 5 . Note that this range includes 0. In other words, the data wander so much, and therefore our estimate of the slope wanders so much, that it probably wanders over 0. And if the slope is 0, then there is no relationship; we cannot confidently reject the possibility that there is no relationship between visits to the mayor and illness. Hence, one of the most important implications of the standard errors is whether we can be 95% confident that the slope coefficient is not 0 (no relationship between X and Y). A quick rule of thumb here is that if the coefficient is greater than twice the standard error, then it is significant at the 95% confidence level.

There are four ways of reporting this information, all of which can be found in the computer readout (see Table 57.1). The first way is simply to report the standard errors and let the reader do his or her own multiplication or division by 2 (technically 1.96). The second way is to report the results of the division. That is, report the ratio between the coefficient and the standard errors. Remember that we want 2 or more standard deviations away from 0 in order to be confident in rejecting the possibility of no relationship.

So we want a ratio greater than 2 (technically 1.96). This is known as a *t* score, *t* ratio, or *t* statistic. Third, you could report the *p* value, which is the probability of the coefficient being zero. Finally, one could report the confidence interval itself, though this is rarely done.

Since these measures are just different ways of reporting the same data, authors and journals vary in the formats they favor. Furthermore, when reporting regression results in a table, authors usually highlight statistically significant findings with asterisks. That is, coefficients that the author is 95% confident are not 0 might receive a single asterisk next to the standard error; coefficients with 99% confidence get two asterisks, and so on. For example, a coefficient of 3 with a standard error of 1.5 would have a *t* statistic of 2, or a *p* value of .05, and regardless of which of these were reported, it would receive an asterisk next to it.

Multivariate Regressions

Bivariate regressions are useful, but usually when we want to explain something, we have more than one independent variable that we want to control for. Let's go back to our flu example. What if we finally realize that the number of Norwegians in a zip code affects how many people there get the flu, or maybe we want to control for whether the district gave out flu shots or access to health care. In physics and engineering, when you start adding more variables, things start getting really complicated, as do the mathematics to explain them. Good news: That's not true in statistics. Regressions work almost exactly the same way with 2 variables as with 3, 4, or 100. Indeed, the real payoff of regression analysis comes when we move from the bivariate case of *X* causing *Y* to the multivariate case of two or more different *X*'s causing *Y*.

Why? Put simply, the coefficients produced by multivariate OLS tell us the amount that *Y* changes for each unit increase in each *X*, holding each of the other independent variables constant. This ability to hold constant is hugely important. In controlled laboratory experiments, scientists hold everything constant except for one causal variable. They then vary that one causal variable and see what effect it has on the phenomenon they are studying (the dependent variable). It is their ability to isolate all causal factors except one that gives controlled experiments their great explanatory power.

Multivariate OLS allows us to hold variables constant mathematically. Say we want to study the effect of two independent variables (X_1 and X_2) on a dependent variable (*Y*). When calculating the effect of a change in X_1 on the average *Y*, OLS mathematically partials out the effect of X_2 on *Y*. It also mathematically strips out the parts of X_2 that might affect X_1 . This has the effect of isolating X_1 and its effect on *Y*. Hence, we can interpret the coefficient on X_1 as being the average change in *Y* for a unit change in X_1 ,

holding all other variables constant (or controlling for all other variables). OLS simultaneously does this for each of the other independent variables included in the regression. So the coefficient for each independent variable can be interpreted as the effect of that variable, holding all of the others constant. This is fantastic for political scientists since it is usually not practical or possible for us to conduct laboratory experiments. It means that in situations where we cannot exert experimental control to produce data and thereby test hypotheses, we can instead use OLS to exert statistical control over data we collect and thereby test hypotheses.⁴

In the multivariate case, we still have just one *Y* (just one dependent variable, just one effect we are trying to explain), but now we can estimate the effects of multiple different *X*'s (multiple different causal variables). The concepts all work exactly the same way as for the bivariate case. The basic difference is the number of dimensions. In the bivariate case, we have a two-dimensional plane of data (*x* on the *x*-axis, *y* on the *y*-axis), and OLS fits a one-dimensional line through it. When we increase to the three-variable case (two independent variables and one dependent variable), then we have a three-dimensional cube of data, and OLS fits a two-dimensional plane through it. But we still interpret the coefficients and standard errors in just the same way as in the bivariate case.

For a quick example, let's go back to our mysterious flu case and say that we suspect Norwegians are much more susceptible to this flu than everyone else. So we add to our data set an independent variable that tracks the percentage of each zip code's population that is Norwegian. Now instead of regression of *Y* on *X* (illness on visits to the mayor), we now regress *Y* on X_1 and X_2 (illness on visits to the mayor and %Norwegians). The equation reads like this:

$$\text{illness} = 2.0 (\text{visits}) + 5.8 (\% \text{Norwegians}) + 106. \quad (2)$$

We can interpret this equation in the following manner: If we hold constant the percentage of population that is Norwegian, then for every one additional visit to the mayor, there will be an average of 2.0 more sick people. If we control for the number of mayor visits, then for each additional 1% increase in the Norwegian population, there will be an average of 5.8 more sick people. Finally, if no one visits the mayor and there are no Norwegians, then we should still expect an average of 106 cases of illness.

Gauss-Markov Assumptions

The Gauss-Markov theorem proves that OLS produces the best linear unbiased estimators (i.e., the best fitting lines). But it only works under certain conditions. Indeed, we

have already seen examples of some of the things that can go wrong when using OLS. These are really just examples of Gauss-Markov assumptions that need to be fulfilled in order for OLS to work properly.

Assumption 1: A Continuous Dependent Variable. OLS assumes that your Y is a continuous variable (e.g., population, gross domestic product, or percentage of votes received). OLS does not work when your dependent variable is a category (e.g., Republican, Democrat, Independent, Socialist, Capitalist, or Communist) or dichotomous (e.g., war or peace, win or lose, yes or no). In these cases, the fix is to use a slightly more advanced regression technique, such as probit or logit. But there is no need to worry if any of the independent variables is a noncontinuous variable; OLS can handle that just fine.

Assumption 2: A Linear Relationship. Since OLS draws lines through data points, the relationship you hypothesize must be linear. The flip side is this: Just because OLS fails to produce significant coefficients, it does not mean that there is no relationship between X and Y . Consider a U-shaped relationship between X and Y . If you performed regression analysis on data describing this relationship, the result would be a line with zero slope and a very low R^2 . You would likely walk away from such a regression mistakenly thinking, “No relationship here.” OLS would likewise fail to properly recognize exponential relationships, logarithmic relationships, quadratic relationships, and so on.

What is the proper fix? Transform the data. That is, if you suspect a nonlinear relationship, then perform a mathematical operation on the data that would turn it linear for the purposes of testing. For example, if we suspected the inverse U shape that follows, we might divide the data and invert one half of it and then use OLS to try to fit a line to it, or we might perform one regression on the lower half of the data (looking for a positive slope) and another regression on the higher half (looking for a negative slope). You can likewise use logarithms, exponents, and squares to transform other types of data where appropriate.

Assumption 3: The data are accurate and the sample is random. As with all statistical analysis, the results from OLS are only as good as the data. Hence, OLS assumes that measurement errors are minimal and that existing errors are random. In other words, there should be no systematic bias in the data. So in the preceding mayor–illness regressions, if you had accidentally taken most of your data from heavily Norwegian neighborhoods (or from neighborhoods with no Norwegians at all), then you would have gotten inaccurate regression results because your sample would not have been representative.

Beyond these three fairly obvious assumptions, the way to think about the Gauss-Markov assumptions is to ask what can go wrong with your regressions. Since we care mostly about estimating the slopes correctly, there are usually only two things that can go wrong in OLS:

Either the estimates of the coefficients can be off (biased), or the standard errors can be off (inefficient). Therefore, we should focus on conditions that can cause these problems.

Assumption 4: No Model Specification Error. Put simply, in addition to linearity (Assumption 2 above), this means that all relevant X 's should be included in the model, and irrelevant X 's should not be included in the model. Omitting a relevant variable (called *omitted variable bias*) can result in biased estimation of the coefficients, while including irrelevant variables can inflate the standard errors of the other X 's.

Why? Because if you control for an X that does not really matter but is highly correlated with an X that does, it will steal some of its explanatory power. For example, say a lot of the people who were visiting the mayor and getting sick happened to be Democrats. If you included party affiliation (e.g., Democrat vs. Republican) as one of your regressors, then OLS would look at the data and say, “Wow, visits to the mayor matters a lot . . . and so does being a Democrat.” However, we know that being a Democrat does not make you ill (regardless of how it might make Republicans feel). But if you were to include party affiliation in the regression model, then the coefficient for visits to the mayor would be smaller than it should be because the coefficient for Democrat would steal from it.

Assumption 5: Homoskedastic Errors. Homoskedasticity is Greek for equally spread out. It refers to the fact that OLS requires that the errors all have the same spread (variance) for each value of the independent variable. In the preceding flu example, the visits to the mayor and Norwegians data might err differently for different zip codes. The variance might be quite wide in downtown zip codes, where many people visit city hall daily and others not at all. Meanwhile, out in the suburbs, people might generally visit the mayor one or fewer times per year; thus, when viewed as a group, their individual visits are each closer to the group's mean. But OLS assumes that homoskedasticity. If you have heteroskedastic errors, then OLS will still produce good coefficients, but the standard error estimates will be too small. Therefore, you could wind up mistaking a significant finding for an insignificant one. This is commonly a problem with regressions involving data on multiple geographic areas (countries, states, cities) or organizations (firms, political parties). The solution is to use a slightly modified form of OLS that weights its estimates. In practice, this usually just means entering an additional command in your computer software.

Assumption 6: Errors are normally distributed. Ideally, if one could measure all of the residuals and then plot them on a graph, they should have a Gaussian or so-called normal distribution. This is possibly the least important assumption, but some researchers argue that where it holds true OLS produces the best estimates.

Assumption 7: No Autocorrelation. This means that the residuals should not be correlated with each other

across observations. This is rarely a problem with cross-sectional regressions. Cross-sectional regressions are those that analyze different units (e.g., nations, states, or companies) during a snapshot in time. For example, a regression of economic growth in 100 countries during 2005 is a cross-section. However, if you want to analyze data across time (e.g., economic growth in the United States from 1900 to 2008), known as *time series*, then autocorrelation becomes a problem. Why? Consider presidential popularity, economic growth, or the stock market. Today's value of these variables depends, at least somewhat, on their value yesterday, the day before, and the day before that. Therefore, some of today's errors (wandering data) will be correlated with or caused by yesterday's errors, those of the day before, and so on. This correlation implies that some variable has been left out. OLS does not deal with this well. There are techniques to diagnose autocorrelation, and there are some simple fixes available (such as including year dummies or a time trend variable). But you often have to use a different regression technique, such as time-series analysis (for single units observed over a long period of time, e.g., stock prices) or time-series cross-section (for multiple units observed over a period of time, e.g., cross-national comparisons of economic growth).

Assumption 8: The errors should not correlate with any of the X's. Some statisticians argue that this is the only important assumption. It is actually another way of saying that you have not left any variables out of your equation. How can the error estimates be correlated with any of the X's? The error term actually represents all causal factors not included as individual X's. In the preceding illness example, this would include everything from random factors like the weather and nose picking to possibly important factors like the number of Norwegians. And we know that if you omit a variable and it is correlated with any of the X's, then the regression will produce biased results.

Dummy Variables

Although OLS does not work when the dependent variable is dichotomous or categorical, it can handle dichotomous independent variables just fine. A dichotomous or dummy variable is a variable that can be coded only as 1 or 0. For example, you might use dummies to code variables like political party, ethnicity, country, or war. But the interpretation of dummy variables is very different from that of continuous variables. Specifically, you do not interpret the coefficients of dummy variables as slopes of a line. Instead, you interpret dummies as creating a separate regression line, with the same slope but a different intercept.

Let's see a simple example. Say you want to explain differences in worker salaries ($Y = \text{salary}$). You hypothesize

that salary is a function of education. You also suspect that gender discrimination affects salary decisions; therefore, you want to control for gender too. So your model is $\text{salary} = \text{schooling} + \text{gender}$. Salary and years of schooling are continuous numbers, but how should you handle gender? The solution is to create a dummy, say male, in which you code 1 for men and 0 for women. Note that you do not create a gender dummy since it would not be intuitively clear what a 1 or 0 gender would mean. You also do not create both a male and a female dummy since this would be redundant (knowing the value for male makes a female dummy unnecessary); this would also crash the mathematical solution to the regression, but this chapter leaves that part of the explanation to your statistics class.

Next, you would collect salary, schooling, and gender data on a large number of individuals. You would then run a regression on the data and observe the coefficients and standard errors just like with any other regression. But the interpretation of these coefficients is a bit different for the male dummy. Let's say the results produce a regression line that looks like this:

$$\text{salary} = \$5,500 (\text{schooling}) + \$3,000 (\text{male}) + \$9,000. \quad (3)$$

These results suggest that workers make \$5,500 more in income for each additional year of schooling they receive. It also shows that male workers make \$3,000 more than females. In other words, the male dummy can be 1 or 0, while its coefficient is \$3,000. Together, they multiply to be either 0 (for females) or \$3,000 (for males); hence, dummies simply add to the intercept term. In other words, the model implies two different regression lines, one for females and one for males. (See Figure 57.6.) The slopes are the same, but the intercepts are different, with the difference between the two lines being the dummy's coefficient.

$$\text{male salaries} = \$5,500 (\text{schooling}) + \$12,000. \quad (4)$$

$$\text{female salaries} = \$5,500 (\text{schooling}) + \$9,000. \quad (5)$$

Another way of interpreting this regression is that since you always include one fewer dummy than the number of categories, the coefficients for the dummies tell how much effect the variable has relative to the missing category. This can be seen more clearly if we add race to Equation 5. Race is a categorical variable but not a dummy variable because there are more than two race categories. We therefore turn race into dummy variables by creating a dummy for each category we want to analyze:

$$\text{income} = \text{schooling} + \text{female} + \text{Asian} + \text{black} + \text{Latino}. \quad (6)$$

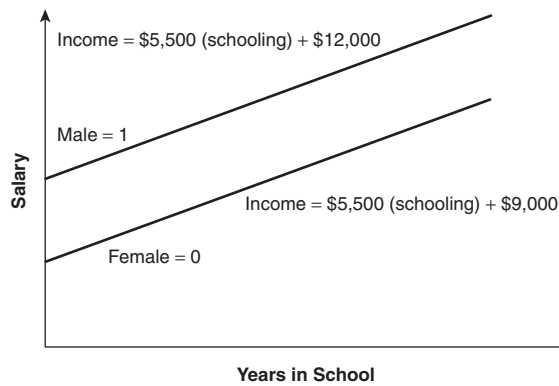


Figure 57.6 Dummy Variables

In this case, we suspect that racial discrimination also affects salaries. Let's assume that our hypothesis focuses on four major race categories (Asian, black, Latino, white). To control for race, we would include in the regression model dummies for only *three* of these categories, say Asian, black, and Latino. But we do *not* put in a dummy for white because it would be redundant. The resulting coefficients for Asian, black, and Latino would therefore tell how much workers in these race categories earn relative to workers in the missing category, white. Let's say the regression results look like this:

$$\text{salary} = \$3,723 (\text{years in school}) + \$509 (\text{male}) - \$680 (\text{Asian}) + \$1,920 (\text{black}) - \$900 (\text{Latino}) + \$15,680. \quad (7)$$

These hypothetical regression results suggest the following, at least for the sample for which we collected data:

1. Each year of schooling results in a salary increase of \$3,723 (holding race and gender constant).
2. Male workers make \$509 more than female workers (holding race and schooling constant).
3. Asians make \$680 less than whites (holding gender and schooling constant).
4. Blacks make \$1,920 more than whites (holding gender and schooling constant).
5. Latinos make \$900 less than whites (holding gender and schooling constant).
6. The salary of an uneducated, white female worker (i.e., a value of 0 for all variables) is \$15,680.

Interaction Terms

A final device often used in social science regressions is an interaction, or multiplicative, term. An interaction

term multiplies two independent variables together. They are used to model hypotheses in which the effect of X_1 on Y is conditional on X_2 (or vice versa). Let's see how this might work.

Sticking with our preceding example, say we hypothesize that salaries are a function of education and experience. But we believe that the effects of education on salary are conditional on experience. That is, the effect on salary of a worker's education will depend on his or her experience on the job. An experienced worker will get a larger salary bump from an MBA than someone fresh out of college. Such a regression model would look like this:

$$\text{salary} = X_1 (\text{education}) + X_2 (\text{experience}) + X_3 (\text{education*experience}), \quad (8)$$

where (education*experience) is the interaction term.

The interaction term's coefficient, X_3 , tells us the effect on income of being educated and experienced that is not explained by education and experience when they are considered separately. More precisely, X_3 tells you how much the effect of education (on salary) changes per unit increase in experience (and vice versa). It therefore tells you how the effect of education (on salary) is conditional on experience (and vice versa). The term *vice versa* gets repeated here because the two statements are mathematically equivalent. Statistics cannot tell you which independent variable drives the other in the interaction term. Rather, good theory should come first and inform us how to interpret the interaction term.

Notice that X_1 no longer tells you the average effect of a unit increase of schooling on income, holding experience (X_2) constant. Instead, the coefficient on education now tells you the effect of 1 year of education when experience = 0.

Likewise, X_2 no longer tells you the average effect of a unit increase of experience on income, holding education (X_1) constant. The coefficient on experience now tells us the effect of 1 year of experience when education = 0. This can be seen a lot easier in an example. Say we run the preceding regression on salary, education, and experience data and produce the following coefficients:

$$\text{salary} = \$3,000 (\text{education}) + \$700 (\text{experience}) + \$250 (\text{education*experience}) + \$5,000. \quad (9)$$

Therefore, if we start with a totally uneducated worker and add education 1 year at a time, we would begin to get the results in Table 57.2.

The regression results tell us that an uneducated worker with no experience could expect a salary of \$5,000. Every year of additional education would add not only \$3,000 in salary directly from schooling but

<i>Education (in Years)</i>	<i>Salary (\$)</i>
0	$0 + 700 (\text{experience}) + 0 (\text{experience}) + 5,000$
1	$3,000 + 700 (\text{experience}) + 250 (\text{experience}) + 5,000$
2	$6,000 + 700 (\text{experience}) + 500 (\text{experience}) + 5,000$

Table 57.2 Interpreting Interaction Terms

also some additional income that would depend on the amount of experience (i.e., interaction term). Hence, the \$3,000 coefficient tells us the effect of 1 year of education when experience = 0. The \$250 coefficient tells us how much the effect of education (on income) changes per unit increase in experience (or vice versa). Note that this means that the effect of education on income is different for different levels of experience (and vice versa). Workers with only 1 year of education can expect each year of experience to add \$700 + \$250 to their incomes. But workers with 2 years of education will get more out of their experience: \$700 + \$500. Therefore, the effects of education are conditional on the amount of experience (and vice versa).

When including an interaction term, many researchers also include its components separately, as was done here. This is because they want to show that the interaction term is significant even after they control for its components. However, there is nothing wrong with a theory that hypothesizes that the interaction term alone is what matters.

Future Directions

The English poet Alexander Pope once wrote that “a little learning is a dangerous thing.” So consider yourself warned: You have now learned enough about regressions to be dangerous. But in order to be useful, you need to learn more. The good news is that if you have taken the time to understand the basic concepts described in this chapter, then learning more should be easy. In fact, SAGE Publications offers a special series, *Quantitative Applications in the Social Sciences*, of more than 160 little green books, each dedicated to a different statistics topic. They are generally very well written, highly accessible, light on math and theory, but heavy on examples and applications. Therefore, if you mostly understood this chapter, but want to nail down some individual concepts better, or dig a bit deeper, then these books should be your next step. Some of the most relevant booklets have been listed in the references,

along with some very useful textbooks, articles, and book chapters.

Notes

1. Probit, logit, and scobit are used when the dependent variable is not continuous. Distributed lag models are used to analyze variables that change over time and where the current value of the dependent variable is partly explained by its previous (“lagged”) values. Panel regressions are used to analyze variables that change across both time and space (e.g., country, state, city).
2. Good researchers also include an explanation of the causal mechanism (precisely how X causes Y to vary) in their theories.
3. Many spreadsheet programs, such as Excel and CALC, can perform basic regressions as well.
4. Readers who need more convincing, and want to see exactly how OLS exerts mathematical control over data, should consult the books recommended at the end of this chapter.

References and Further Readings

- Achen, C. (1982). *Interpreting and using regression*. Beverly Hills, CA: Sage.
- Achen, C. (1990). What does “explained variance” explain? *Political Analysis*, 2, 173–184.
- Beck, N. (2008). Time series cross section methods. In J. M. Box Steffensmeier, H. E. Brady, & D. Collier (Eds.), *The Oxford handbook of political methodology* (pp. 475–493). New York: Oxford University Press.
- Berry, W. D. (1993). *Understanding regression assumptions*. Newbury Park, CA: Sage.
- Berry, W. D., & Feldman, S. (1985). *Multiple regression in practice*. Beverly Hills, CA: Sage.
- Berry, W. D., & Sanders, M. (2000). *Understanding multivariate research: A primer for beginning social scientists*. Boulder, CO: Westview Press.
- Brambor, T., & Clark, W. R. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis*, 14, 63–82.
- Chatterjee, S., & Wiseman, F. (1983). Use of regression diagnostics in political science research. *American Journal of Political Science*, 27, 601–613.

- Fox, J. (1991). *Regression diagnostics: An introduction*. Newbury Park, CA: Sage.
- Greene, W. H. (2000). *Econometric analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- Gujarati, D. N. (2009). *Basic econometrics*. New York: McGraw Hill.
- Hardy, M. A. (1993). *Regression with dummy variables*. Newbury Park, CA: Sage.
- Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression*. Thousand Oaks, CA: Sage.
- Kennedy, P. (2003). *A guide to econometrics* (5th ed.). Cambridge: MIT Press.
- King, G. (1986). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, 30, 666–687.
- Lewis Beck, M. S. (1980). *Applied regression: An introduction*. Beverly Hills, CA: Sage.
- Lewis Beck, M. S., & Skalaban, A. (1990). The *R* squared: Some straight talk. *Political Analysis*, 2, 153–171.
- Pevenhouse, J. C., & Brozek, J. D. (2008). Time series analysis. In J. M. Box, Steffensmeier, H. E. Brady, & D. Collier (Eds.), *The Oxford handbook of political methodology* (pp. 456–474). New York: Oxford University Press.
- Pollock, P. H. (2006). *A STATA companion to political analysis*. Washington, DC: CQ Press.
- Pollock, P. H. (2008). *An SPSS companion to political analysis*. Washington, DC: CQ Press.
- Schroeder, L., Sjoquist, D. L., & Stephan, P. E. (1986). *Understanding regression analysis: An introductory guide*. Beverly Hills, CA: Sage.
- Williams, F., & Monge, P. R. (2000). *Reasoning with statistics: How to read quantitative research*. Orlando, FL: Harcourt Brace.
- Wooldridge, J. M. (2008). *Introductory econometrics: A modern approach* (4th ed.). Mason, OH: South Western.

58

CONTENT ANALYSIS

MARIJKE BREUNING

University of North Texas

Content analysis is, as its name suggests, the analysis of the content of communications. Researchers use content analysis to make statements about the meaning, impact, or producers of those communications. Depending on the purpose of the specific research project, analysts may focus on the literal content or seek to extract deeper (or latent) meanings.

This multiplicity of purposes has led content analysts to use a variety of strategies for analyzing text systematically. Some of these strategies, such as word counts, are easy to replicate, whereas other forms are far more interpretive and dependent on the judgment of the individual who codes the text. Most forms of content analysis yield quantitative indicators. Indeed, some would define quantification as an essential aspect of content analysis (e.g., Weber, 1990). Others view it as preferable but not essential (Berelson, 1952; Holsti, 1969).

Content analysis is not new. According to Krippendorff (1980), empirical studies of communications can be dated back to the 1600s. More immediate ancestors to modern content analysis, however, are studies that sought to evaluate the content of mass media in the early 20th century and Nazi propaganda during World War II (Berelson, 1952; Krippendorff, 1980). As a method for studying communications, content analysis has been an especially popular methodology in the field of (mass) communication.

Holsti (1969) reported a trend toward a more frequent use of content analysis, as well as its application to a broader

array of problems, including subjects of interest to political scientists. He furthermore noted an emerging tendency for content analysis to be used in combination with other social science research methods and a move toward computer-assisted content analysis.

This chapter emphasizes quantification, although it also discusses some of the trade-offs between quantitative and qualitative forms of content analysis. After discussing definitions and forms of content analysis, the chapter describes some of the issues in designing content analysis studies, with the objective of giving the reader the basic tools for evaluating whether this methodology might be useful in her or his research. The chapter then turns to the issue of reliability and stability, which are of particular importance when using human coders, and subsequently turns to a discussion of the emerging strategy of using computer assistance in the coding of text for content analysis. The chapter ends with an assessment of the future of content analysis in political science.

What Is Content Analysis?

According to Weber (1990), content analysis is a “research method that uses a set of procedures to make valid inferences from text” (p. 9). This concise definition captures the essence of content analysis very well, although it may be worth adding that text is not the only content that might be

subjected to analysis. (Transcripts of) oral communications, as well as visual communications, could also be subjected to this type of analysis. This chapter, however, limits its scope to the content analysis of text (or at least verbal material) and does not consider the analysis of visual communications.

Beyond making valid inferences from text, most content analysis “seeks to quantify content in terms of predetermined categories and in a systematic and replicable manner” (Bryman, 2004, p. 181; see also Holsti, 1969). In other words, content analysis endeavors to analyze text in a systematic, empirical manner that is made sufficiently explicit to permit replication. Generally, this means that content analysis proceeds on the basis of instructions that enumerate explicit categories. Consequently, Babbie (2004) has described content analysis as “essentially a coding operation” (p. 318). Although this is accurate, it also sells content analysis short as a method for analyzing the content of communications.

The coding operation is at the heart of content analysis, but content analysis cannot be reduced to coding, just as public opinion research cannot be reduced to the survey instruments often used to ascertain public opinion. The coding of text or other communications permits the analyst to ascertain patterns and test hypotheses about those communications. Holsti (1969) maintained that content analysis “must be undertaken for some theoretical reason” (p. 14). Although not all studies employing content analysis satisfy that condition, such studies are generally undertaken to answer some question that is either of scientific interest or of political (or professional) relevance. In other words, studies may employ content analysis for a variety of purposes. Holsti summarized these purposes into three groupings:

1. Content analysis may be used to describe characteristics of communications. For instance, a researcher may wish to discern trends in the content of newspaper or other media outlets or analyze the rhetorical style of a decision maker. An example of this type of research in political science is Breuning, Bredehoft, and Walton’s (2005) analysis of academic journal content.
2. Content analysis may be used to infer psychological or other characteristics of the speaker. For instance, Hermann’s (1980, 2002) *Leadership Trait Analysis* is based on content analysis. She used a coding scheme that is informed by psychological theories to analyze the (spontaneous) remarks of political leaders in order to make assessments about various personality traits.
3. It is further possible to use content analysis to assess the (potential) impact of communications. Eshbaugh Soha’s (2006) study of political impact of presidential speeches is an example of this type of analysis.

Within political science, content analysis has been used for all three of these purposes, although not with the same frequency. Studies that use content analysis to evaluate leader personality, motivation, or both, tend to be more

plentiful in political science than studies of the impact of communication. Studies that describe the characteristics of communications tend to be more plentiful in the study of (mass) communication than in political science.

A distinct benefit of content analysis is that it is an unobtrusive research method (Babbie, 2004). The advantage of such a research method is that, one, it does not require the cooperation of the subject under investigation, and two, the subject will not alter her or his behavior as a result of awareness of being tested. The second point is important. There is evidence, for instance, that survey respondents on occasion provide socially acceptable answers rather than truthfully reporting their behavior. They may say they voted, because they think they should have, when in fact they stayed home. The first point is relevant to the study of political decision makers and foreign policy decision making. Although it can be useful to understand what motivates decision makers, it is highly unlikely that such individuals would make themselves available for psychological testing. Some researchers (e.g., Hermann, 1980, 2002; Walker, Schafer, & Young, 1998; Winter, 2005) have therefore devised research strategies that rely on content analysis of remarks and speeches in order to evaluate decision makers’ personalities and motivations. It is in this area of political science that content analysis is used most consistently.

Other benefits of content analysis include that it is relatively easy to undertake. It requires no special equipment or access to significant research funds (Babbie, 2004). Whereas survey research can be expensive, a study using content analysis can be completed for very little money. A single investigator with access to the relevant textual material for coding can complete a content analysis study, although a very large content analysis-based study may require multiple human coders or access to content analysis software in order to complete the study in a timely fashion. Paying human coders or purchasing content analysis software would, obviously, add to the expense of implementing a content analysis study.

Further, an investigator can much more easily repeat a portion of the study than would be the case with survey research (Babbie, 2004; Bryman, 2004). This includes the determination that an additional dimension needs to be considered in the analysis. It requires going back through the text to code the additional variable, which can be time-consuming when human coders are used, but it remains feasible as long as the text remains available. Repeating the analysis or adding another dimension to be coded is quite easy when a computer-assisted content analysis strategy is used.

Finally, content analysis lends itself to studying trends over long stretches of time (Bryman, 2004). For instance, Eshbaugh-Soha (2006) analyzed presidential speeches across a 50-year period.

All of these advantages make content analysis useful and attractive. On the other hand, the method is limited to

the investigation of text and recorded human communications (Babbie, 2004; Bryman, 2004).

Forms of Content Analysis

As already mentioned, content analysis generally refers to quantitative assessment of various aspects of text. This chapter places its emphasis on such analysis. As outlined previously, content analysis is defined by the quest to analyze text in a manner that is systematic, valid, and replicable. The first and last criteria are most easily satisfied through quantification. Although qualitative content analysis may be equally valid in its assessments of text, it is much less likely to be systematic, and it is exceedingly difficult to replicate. Although quantification can have important limitations, it has the distinct advantage of transparency: An explicitly formulated research design for a systematic and quantitative research design not only can be replicated but also allows any reader to evaluate how the investigator arrived at her or his conclusions.

Qualitative analysis does not inherently require such explicit research design and quite often depends on the expertise of the investigator (see, e.g., Neumann, 2008). The advantage of such analysis is that an investigator with substantive expertise may be able to identify nuances that a quantitative analysis misses. On the other hand, qualitative analysis, because it often lacks the sort of explicit coding scheme that is required for quantitative analysis, provides much less of a hedge against investigator bias. This led Babbie (2004) to suggest that an investigator engaged in qualitative content analysis must carefully search for disconfirming evidence—and report meticulously on any elements in the text that are inconsistent with the expected findings—to guard against a tendency to focus on the elements of the textual material that confirm the investigator's expectations. If this is not done, there is the distinct risk that qualitative content analysis leads an investigator to confirm her or his expectations when these are not in fact supported by the evidence.

Quantitative content analysis, on the other hand, is more explicitly systematic and less dependent on the ability of an investigator to counteract the tendency to focus on confirming evidence. However, it may be less sensitive to contextual and cultural cues. In addition, the validity of quantitative content analysis depends on the adequacy of the coding scheme employed to analyze the text.

First, it is far easier to code manifest than latent content. Manifest content refers to the surface meaning of text, whereas latent content refers to the deeper or symbolic meaning. Content analysis has generally favored a focus on manifest content. Indeed, it has often been defined in terms of explicit and systematic coding rules (see, e.g., Holsti, 1969). In contrast, the analysis of the deeper, underlying meaning of text has more often been the province of those who favor qualitative content analysis or

also discourse analysis (e.g., Neumann, 2008). As discussed previously, the latter type of analysis is less transparent, and replication is difficult or impossible. The focus on explicit and systematic coding rules does not mean that content analysis avoids interpretation but rather that it separates the data-gathering operation (the coding, counting, or both) from the interpretation of the results.

Second, the coding scheme needs to be carefully designed not only to ensure that it is explicitly stated and replicable, but also to ensure that it is grounded in the research question. In other words, the categories that are employed should be theoretically justified so that the resulting data help the investigator draw valid inferences from the text. What constitutes an appropriate coding scheme will depend on the research question the investigator seeks to test.

The distinction between manifest and latent content is perhaps somewhat artificial. Frequently, coding schemes that are based on relatively straightforward elements, such as word counts, do in fact seek to evaluate some aspect of latent content. In such cases, the specific words that are counted may have been chosen to reveal latent content.

For instance, in an analysis of the rhetoric of two ethnic nationalist Flemish parties in Belgium, Breuning and Ishiyama (1998) counted the references in the party platforms to the terms *foreigners* and *immigrants*. They theorized that the latter term has the connotation that the speaker or author perceives the persons who are being described as individuals who will integrate into the society and that the former describes individuals who are perceived as (permanent) outsiders. They further theorized that the extent to which a party preferred to use the term *foreigner* rather than *immigrant* when discussing nonnative populations was indicative of the party's xenophobia. In other words, Breuning and Ishiyama used word counts of specific terms to measure the differences in xenophobic attitude between these two ethnic nationalist parties. The coding scheme was transparent and easy to replicate but was designed to evaluate latent content.

The advantage of using carefully chosen but explicit coding schemes to reveal some aspect of latent content is that they are transparent and permit the researcher to show exactly how she or he arrived at the study's interpretations and conclusions.

In addition to simple word counts, content analysis coding schemes may also record valences—that is, the value associated with the word that is coded. It does, after all, make a difference whether a word is used in a positive or negative construction. The automated coding scheme developed to study decision-makers' operational codes includes valence measures (Walker et al., 1998).

Another variation is the use of thematic coding. In this case, the investigator is looking for evidence of specific themes in units of text. Theme coding can be tricky, since it involves judgment rather than simple counts. However, there are situations where thematic coding is theoretically

justified—and word counts may simply not provide the investigator with an appropriate measure. In the creation of a coding scheme for thematic coding, extra care must be taken to make very clear and explicit the criteria for judging a unit of text to belong in one versus another category. Such a coding scheme may require pilot testing with several different coders to ensure that these coders share a common understanding of the coding instructions. The mechanics of designing a content analysis scheme are explored further in the next section of this chapter.

Issues in Designing Content Analysis Studies

It can be difficult to devise a good content analysis study, although this also depends on the objective of the study. If the objective is to produce a study that is descriptive of some aspect of text, the construction of a coding scheme may not need to be overly complicated. In such a case, manifest content is coded to make systematic observations about that manifest content. At times, such studies have been accused of being atheoretical (Bryman, 2004). Whether this is a problem depends on the contribution a specific study seeks to make. For instance, Breuning et al. (2005) sought to analyze the content of academic journals and did not seek to test any theory. They simply sought to demonstrate in a quantifiable way who and what got published in the journals that they investigated. Such data provide insight into, for instance, the sort of work that is published in specific journals.

If, on the other hand, the study seeks to make inferences about a speaker or author or about the text's impact, it can be much more difficult to determine what, exactly, needs to be coded. In that case, manifest content is coded in order to make inferences about underlying meaning or latent content. Content analysis studies that are undertaken for such objectives need to be thoroughly grounded in relevant theories, which can help to justify the validity of the measures. For instance, Hermann (1980, 2002) and Winter (2005) devised their content analysis schemes on the basis of psychological theories. Both sought to make inferences about the personality traits of decision makers on the basis of their interview responses and speeches.

How should an investigator go about designing a content analysis study, irrespective of whether she or he seeks to focus on manifest or latent content? There is not one single, correct way to design a content analysis study. The investigator needs to carefully consider a variety of issues in order to design a content analysis that helps her or him find answers to a specific research question. There are several sources that can provide guidance. Hermann (2002) presented a set of eight questions that are designed to help a researcher determine whether content analysis is an appropriate research strategy. Holsti (1969), Krippendorff (1980), and Weber (1990) each discussed

many of the trade-offs and mechanical details of designing content analysis studies. All of these sources are useful for investigators who would like more detailed guidance in designing a content analysis study. What follows here is a condensed guide to the sort of issues that investigators need to consider in designing such studies. It is intended to enable the reader to evaluate whether this methodology might be appropriate in her or his own research projects.

The first step is to devise a research question and ask whether content analysis would be an appropriate research strategy. If the question can be answered using text or other communications, then a content analysis may well be a useful strategy. It is important to check at this early stage whether a sufficient volume of appropriate text is available and accessible and in what form it is accessible. If one plans to conduct a computer-assisted content analysis, it will be extremely useful if the materials are available in electronic format. Contemporary text is far more likely to be accessible in such a format than are older documents. This does not make it impossible to conduct a computer-assisted content analysis of older documents, but doing so would require converting such documents to a machine-readable electronic format. That is an extra step that might be time-consuming to complete.

The second step involves decisions about the type of analysis to conduct. Content analysis is about generating data to answer a specific research question. Given one's research question, is a quantitative or a qualitative approach appropriate? Is it possible to count word frequencies? Or would a thematic coding scheme be more appropriate? It may initially seem easier to devise a thematic coding scheme, but remember that the data generated in this way are usually less easily replicated than word counts.

It is likely to be more difficult to construct vocabularies for a content analysis study that counts relevant words, because the specific words must be carefully chosen for what they can reveal about underlying meaning. Despite the greater effort that needs to go into designing such a coding scheme, it may be well worth the effort to do so. Word counts that are theoretically informed (and designed to capture the latent meaning of text) derive their validity from that theoretical basis. Furthermore, the explicit nature of word count coding schemes makes them easier to replicate as well as easier to complete through computer-assisted coding strategies.

A distinct advantage of using computer-assisted coding strategies is that reliability becomes a nonissue: The computer finds every instance of the words it is programmed to count, so there is no error. Human coders get fatigued, and their attention wanders, leading them to make mistakes.

Although theme coding can be an appropriate strategy, it is far more difficult to have a computer assist in the task. Theme coding therefore tends to involve human coders, who may perform the coding task in inconsistent ways.

Just as in producing words counts, human coders who are looking for thematic content may get fatigued and make errors as a result. In addition, thematic coding involves a judgment of the presence or absence of a specific theme in a unit of text. Depending on how explicit the coding instructions are and how well trained the human coders are, such judgments may show variation across different human coders.

In both word count and thematic coding, therefore, it is useful to evaluate the quality of the coding operation by taking two measures: One measures the consistency of each human coder, and other measures the congruity of the decisions made by different human coders. The next section delves into these measures in more detail.

The third step involves creating the coding manual and coding schedule (Bryman, 2004). The former is a detailed and explicit set of instructions to coders. It explains what the unit of analysis is and includes all the possible categories for each dimension that will be coded. For word counts, it will list all the words that will be counted as being indicative of a specific dimension, including synonyms and variations of each word. For thematic coding, the coding manual should not only describe the categories, but also provide an example to help coders understand how to evaluate the text they will be coding.

It is extremely important that the categories for each dimension that will be coded be mutually exclusive and collectively exhaustive. This means that nothing should fit into more than one category simultaneously, and everything should fit into one of the categories. There should never be words, phrases, or themes that are part of a dimension that is being coded that cannot find a home in one of the categories. A clear and explicit coding manual is an important key to a successful content analysis study, especially if multiple coders work on the project, but it is also important to persuade others of the validity of the study.

The coding schedule is simply the form that the coder uses to record her or his coding decisions. Coding schedules may be paper, but it is also possible to use electronic spreadsheets or statistical analysis software instead. In the latter instance, the coder bypasses the need for data entry after completing the coding task.

A well-designed manual not only makes the coding task easier and the results more convincing but also ensures that the study is replicable. It is worthwhile to invest time and effort in the creation of a coding manual, pretesting it on a sample of text and revising it—multiple times if needed—to improve the coding categories or the clarity of the instructions, or both. In the process, it is useful to get feedback from others on the coding manual. If human coders will be completing much of the coding task, it also helps to have them each do a small pretest and check not only whether each coder implements the coding scheme as intended but also whether the different coders make substantially similar decisions.

As you ponder the various aspects of the design of a study, keep in mind whether the proposed analysis does indeed capture the goal of the study. Refer back to the research question and ask whether the coding manual will yield data that can reasonably be expected to shed light on that question.

A further issue to consider in designing the study is whether the text that will be analyzed is representational or instrumental (Hermann, 2002). The former type of text can be assumed to faithfully represent (aspects of) the personality, thoughts, or both, of the creator of that text. The latter type of text is generated for an instrumental reason, such as to persuade an audience, and may not reveal much about the speaker or author of that text. Hermann (1980, 2002) favored spontaneous remarks of decision makers for her content analyses, because her objective was to ascertain personality traits, and she expected spontaneous remarks to be far more representational than prepared speeches. Eshbaugh-Soha (2006), on the other hand, was interested in the instrumental use of language by presidents. For him, the prepared speeches of these decision makers were the more appropriate text. In other words, the text chosen for analysis should match the purposes of that analysis, just as the coding manual should be geared to the analytic purposes of the study.

Last, Babbie (2004) made a useful distinction between the unit of analysis and the unit of observation. If, for instance, you are interested in ascertaining the personality traits of a decision maker, then that decision maker is the unit of analysis—the entity about which you want to be able to make statements as a result of your investigation. To arrive at your conclusions about the individual decision maker, however, you may need to analyze multiple (and often a relatively large number) of spontaneous interview responses. Each of these interview responses is a unit of observation. Although a single interview response can be revealing, it is impossible to know whether it represents a typical statement of the subject under investigation unless it is compared with additional responses. Coding multiple units of observation permits the investigator to discern patterns and to make generalizable statements about the unit of analysis.

Human Coders, Reliability, and Stability

The previous section suggests that the use of human coders requires an assessment of the consistency of the coding decisions made by each human coder, as well as an evaluation of the congruity of the decisions made by different human coders. The former is generally referred to as the stability of the coding decisions, and the latter measures the reliability of the coding scheme.

Reliability requires that the coding scheme should lead different human coders to code the same text in the same way. That is, if the coding categories are sufficiently clear

and explicit, trained human coders should exhibit little variation in the way they evaluate text. Whether human coders do indeed exhibit such agreement can be evaluated empirically. In doing so, it is important to compare the agreement on specific coding decisions rather than in the aggregate. For instance, when comparing the overall results of two coders, the data may look similar, but this similarity may evaporate when considering the individual coding decisions. If each coder made 100 coding decisions on one dimension and each assigned 50 of these to Category A, 30 to Category B, and the remaining 20 to Category C, their results look identical. However, when comparing Decision 1, Decision 2, Decision 3, and so on, it may become apparent that this overall congruity hides substantially different judgments regarding the discrete coding decisions. It is therefore recommended to analyze the similarity between individual coding decisions.

In addition, the calculation of an intercoder reliability score must take into account the role of chance. The coders must do better than agreement that results from chance in order for the similarity in their coding decisions to be meaningfully attributed to the clarity of the coding instructions. Both Krippendorff (1980) and Holsti (1969) provide detailed guidance for calculating intercoder reliability.

Well-designed and carefully implemented content analysis studies usually report an intercoder reliability statistic to demonstrate that the coding instrument is indeed sufficiently explicit to permit a high level of agreement between the coding decisions made by different human coders.

In addition, content analysis studies are often concerned with stability, which is defined as the consistency of the coder's decisions across the texts she or he evaluates. Inconsistent decisions can result from coder fatigue but also from slight shifts in the implementation of a coding scheme between the first and last items coded. To evaluate the consistency of the coding decisions across texts and time, coders are sometimes asked to recode some of the text coded early in the project. The coding decisions made by a single coder at these two different times are then compared. If the individual coder makes largely the same decisions, the coding is judged to be stable. The calculation of the stability of the coding can be done using the same statistical tools used to determine the intercoder reliability score. The only difference is that stability measures the consistency of a single human coder, whereas the reliability score measures the consistency of the coding decisions across different human coders.

Computer-Assisted Content Analysis

Content analysis predates the invention of computers. However, the increasing availability and current ubiquity of computers has revolutionized this methodology. Studies that would have required countless hours of meticulous

attention to detail by several human coders can now be completed with much greater reliability and at much higher speed by computers.

Numerous software packages for content analysis have been developed. Some of these have been designed for very specific purposes (such as Diction), whereas other programs try to accommodate a variety of content analysis purposes (e.g., Atlas.ti, ProfilerPlus, and Wordstat). New programs enter the market on a regular basis, and older ones disappear. For this reason, this chapter does not include an overview of available programs.

When investigating software for content analysis, carefully evaluate whether the program suits the purposes for which it will be used. Each program has been developed with a particular purpose in mind and will excel at certain things and be less useful for other purposes. In other words, it is impossible to judge the merits of content analysis software in the abstract and without reference to the purposes for which it is being considered.

The speed with which computers can analyze large volumes of text is not the only advantage. Reliability essentially becomes a nonissue, because a computer program will analyze the same text in the same way no matter how often one asks it to analyze that text. This led West and Fuller (2001) to state that the "value of computer-assisted content analysis, particularly in terms of reliability, is difficult to overstate" (p. 91).

At the same time, it is important to note that computer-assisted content analysis depends for its validity on the careful design of the coding scheme. This remains the work of the investigator. A reliably executed computer-assisted content analysis cannot be better than the coding scheme it implements. Human coders are more likely to point out the flaws in the logic of the study's design, whereas a computer will complete the coding task without question, even if it makes no substantive sense. Hence, it will remain important to obtain feedback on the draft of the coding scheme and to run pilot tests on small amounts of text to ensure that the coding scheme is well designed and appropriate for its purpose. Computers are extremely useful in the mechanical task of coding and can complete such tasks far more reliably than human coders, but the task of theoretically grounding the content analysis, evaluating whether the coding scheme tests what it claims to test, and deciding whether the coding categories meet the appropriate standards of clarity, explicitness, and validity will remain tasks that only the investigator can competently execute.

Future Directions

This chapter endeavors to provide a primer for the sort of research tasks for which content analysis is suited, as well as to suggest basic problems in the design of content analysis studies. The biggest drawback of content analysis has

traditionally been that it was very time-consuming to complete large-scale studies. The advent of computer-assisted content analysis has given new life to this research methodology, and numerous software packages have been developed for this technique. Some of the initial software was suitable for only very specific coding tasks that mirrored the developer's research purposes. More recently, there has been a greater emphasis on the development of software that can be adapted to a multitude of content analysis research designs. Atlas.ti, ProfilerPlus, and Wordstat are part of this new generation of software for content analysis.

One of the critiques of content analysis has long been that it tends to be atheoretical. That was true of some studies but does not adequately capture the more vexing problems that have plagued content analysis in political science. Content analysis is, in some instances, an end in itself. Studies that seek to analyze communicative content systematically, for instance, are examples of this form of content analysis. However, studies that endeavored to analyze the personality characteristics of leaders or the impact of the message often also stopped at analyzing text, when such studies in fact sought to make statements about something beyond the text. Studies focusing on the traits of leaders, for instance, have long claimed that leader personality influences the types of decisions that such individuals make. It is impossible to demonstrate this convincingly on the basis of the results of content analysis alone.

Recent content analysis studies have begun to use computer-assisted content analysis in combination with other variables. Since less time and effort is spent on the completion of the content analysis itself, researchers are more willing to combine the data derived from that analysis with additional variables. Content analysis has thus become a tool for generating data rather than an end in itself.

As computer-assisted content analysis has increasingly become the norm, this methodology has gained an expanded purpose in political science. Content analysis can now be used as a data-making tool that can yield quantitative indicators of aspects of political life that previously were deemed difficult or impossible to measure. For example, rather than claiming that certain personality traits create a disposition toward certain policy responses, investigators can now evaluate systematically whether this is the case. The results of content analysis can now be employed as variables in models that can more directly test the relationship between the traits of leaders and their actions, as well as establish more explicitly the relationship between speeches and their impact on political decisions.

Although there will remain cases where the analysis of the content of text is an end in itself, the more interesting research frontier for political science lies in using this methodology to systematically generate data. Combining this data with additional measures will permit analysts

to evaluate the propositions that drew them to content analysis in the first place.

References and Further Readings

- Babbie, E. (2004). *The practice of social research* (10th ed.). Belmont, CA: Wadsworth.
- Berelson, B. (1952). *Content analysis in communication research*. Glencoe, IL: Free Press.
- Breuning, M., Bredehoft, J., & Walton, E. (2005). Promise and performance: An evaluation of journals in international relations. *International Studies Perspectives*, 6(4), 447–461.
- Breuning, M., & Ishiyama, J. (1998). The rhetoric of nationalism: Rhetorical strategies of the Volksunie and Vlaams Blok in Belgium, 1991–1995. *Political Communication*, 15(1), 5–26.
- Bryman, A. (2004). *Social science research methods* (2nd ed.). Oxford, UK: Oxford University Press.
- Eshbaugh Soha, M. (2006). *The president's speeches: Beyond "going public."* Boulder, CO: Lynne Rienner.
- Hermann, M. G. (1980). Explaining foreign policy behavior using the personal characteristics of political leaders. *International Studies Quarterly*, 24, 7–46.
- Hermann, M. G. (2002). *Assessing leadership style: A trait analysis*. Hilliard, OH: Social Science Automation. Available at <http://www.socialscience.net/Docs/LTA.pdf>
- Holsti, O. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison Wesley.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.
- Neumann, I. B. (2008). Discourse analysis. In A. Klotz & D. Prakash (Eds.), *Qualitative methods in international relations: A practical guide* (pp. 61–77). Basingstoke, UK: Palgrave Macmillan.
- Walker, S. G., Schafer, M., & Young, M. D. (1998). Systematic procedures for operational code analysis: Measuring and modeling Jimmy Carter's operational code. *International Studies Quarterly*, 42(1), 175–190.
- Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA: Sage.
- West, M. D., & Fuller, L. K. (2001). Toward a typology and theoretical grounding for computer content analysis. In M. D. West (Ed.), *Theory, method, and practice in computer content analysis* (pp. 77–96). Westport, CT: Ablex.
- Winter, D. (2005). Measuring the motives of political actors at a distance. In J. Post (Ed.), *The psychological assessment of political leaders* (pp. 153–177). Ann Arbor: University of Michigan Press.

Software for Content Analysis Cited in This Chapter

- Atlas.ti: www.atlasti.com
 Diction 5.0: www.dictionsoftware.com
 ProfilerPlus: www.socialscienceautomation.com
 Wordstat: www.provalisresearch.com

59

LONGITUDINAL ANALYSIS

SHANE NORDYKE

University of South Dakota

Many of the most important questions that are addressed within the field of political science today involve changes and processes that occur over time. The passing of time is often an integral part of the question being answered. What is the effect of the International Monetary Fund on stability in developing countries? What are the short- and long-term effects of significant policy programs? How do different types of electoral systems affect voter turnout? What variables are important for understanding congressional behavior? How does decision making differ for representatives during election years? These are all important questions within various subfields of political science that in one way or another involve changes that take place over time. As Jennings and Niemi (1978) note, “Questions of persistence and change are also fundamental to an understanding of how people cope with and relate to political phenomena” (p. 333). Longitudinal data allow for analysis that is dynamic in nature and are often better able to address these questions than methods that are available for cross-sectional data. This chapter provides the following:

- A description of longitudinal data
- A brief overview of various methods that are available for its analysis
- A discussion of the benefits and challenges of using this type of data

- The historical development of these methods, including the significant contributions of political science scholars
- A brief description of statistical programs that have been developed specifically for longitudinal data

Longitudinal analysis isn’t a specific methodology or form of analysis. Rather, it refers to a broad category of empirical models and methods that have been developed for evaluating longitudinal data. Studies incorporating longitudinal data have been contributing to the field of knowledge in the social sciences for more than 50 years. Longitudinal data includes any data set that has a group of individual observations repeated at multiple times. A good example is the work by M. Kent Jennings on intra- and intergenerational effects on political attitudes. In 1965, Jennings and Niemi conducted a national probability sample of high school seniors and their parents. Eight years later, they were able to collect data again from more than three fourths of their original sample, providing a second time period of data for his large cross-sectional panel. For the next 40 years, they continued to collect data from this same group and eventually added the children of the original respondents to the sample as well. By taking repeated measures of the same individuals, Jennings (2002) was able to “observe how formative political experiences can affect intragenerational cleavages over the adult life space and how they reflect on intergenerational continuities” (p. 303)

in ways that would not have been possible with a cross-sectional analysis of a group of individuals at only one time.

An analysis of longitudinal data can combine the benefits of cross-sectional analysis in its ability to make between-subject comparisons with the added value of observing changes through repeated observations. The development and expansion of these methods have often occurred in other fields, particularly biomedical sciences, psychology, economics, and education, where it is more common for multiple points of data across time to be collected from a specific identified population. However, it is an important method in the discipline of political science because it allows for dynamic analysis that can address changes over time and allows for more powerful inferences with regard to causality. Recent advances in statistical software programs, which allow for the statistical analysis of longitudinal data sets, and the availability of large-scale longitudinal data sets have made longitudinal analysis more accessible to a wider range of political science scholars.

Cross-Sectional Versus Longitudinal Analysis

Cross-sectional data analysis draws inferences about the population by using a large sample of observations to analyze correlations between variables of interest. Cross-sectional analysis provides a snapshot of the population at a single time. Time-series analysis focuses on analyzing single observations for multiple periods. In time-series analysis, time is the primary, and in some cases the only, independent variable of concern. The analysis of longitudinal data combines the contributions of both by evaluating the same cross section of data at multiple observation points. This is not the same as simply repeating a cross-sectional analysis at a different time; to be considered longitudinal data and to reap the advantages that it provides, the individuals or subjects evaluated must remain fixed over time.

An example can help to illustrate the differences. Suppose that an analyst is interested in understanding voter turnout in presidential elections in the United States. A cross-sectional analysis of voter turnout in a large sample of individual counties across the United States may allow the analyst to draw inferences about the effects of important independent variables such as age, median income, educational level, and party affiliation on voter turnout by analyzing the correlations in these variables between subjects (in this case counties). He or she would not, however, be able to analyze how turnouts have changed over time or how the effect of each of the individual variables changes over time. The most a cross-sectional analysis could say about the effect of age on voting behavior is that voters who are 60 years old this year are more or less likely to vote than those who are 30 years old this year, not how aging itself affects an individual's voting behavior. A time-series analysis, on the other hand, might analyze the

national turnout rate for every presidential election for the past 40 years. In doing so, the analyst discovers changes in the overall turnout rate over time—in year t , voter turnout is 5% greater than in year $t + 4$ —but not how individual voters have changed or what caused the changes in turnout over time. Time-series analysis is also limited in that it can't tell whether 60% of the people vote all the time or if everyone votes 60% of the time. Longitudinal data, analyzed properly, can answer all of these questions by making the most of the information about within-subject differences and between-subject differences. A panel designed to answer this question would likely consist of a large sample of individuals surveyed about their voting habits as well as the independent and control variables that are important to the researcher. The same individuals would again be surveyed repeatedly at intervals appropriate to the research question.

In his article, Lubotsky (2007) is able to expand on previous research analyzing wage changes for immigrants. Previous studies had used averages from cross-sectional data taken from two different age-groups in order to estimate wage changes of immigrants over their first 20 years in the United States without taking into consideration changes in the composition of immigrant populations, which might also affect earnings. By using a longitudinal data set that tracked individuals for this period, Lubotsky was able to overcome analytical weaknesses of the cross-sectional data and find that “the actual earnings growth of immigrants who remain in the United States is considerably slower than that implied by comparisons across decennial censuses” (p. 824).

In addition to analyzing change, longitudinal analysis can also strengthen studies interested in causal relationships. A primary weakness of cross-sectional analyses is that even the most carefully designed studies can speak only to correlation between variables and use those correlations to make guarded inferences about causality. This is true even if the study is repeated several times with different samples:

Whereas the cross sectional method infers causation from the presence of correlations, the longitudinal method permits the use of the far more dependable technique for inferring causation by watching the changes as the specified variables interact over a period of time. The longitudinal method can provide the materials for studying change because it obtains information for each individual over a period of time, in a sufficient and properly selected group, which is then combined with the information for other individuals into common classifications and appropriate summaries. The emphasis is placed on each individual history. (Goldfarb, 1960, p. 8)

With cross-sectional analysis, the strength of the results lies in the correlations between the independent and dependent variables. However, the analyst must constantly be on the lookout for spurious relationships: correlations that appear to exist between two variables that are unrelated but instead are both related to some third variable not included in the model. Even if there is a relationship

between two variables, the nature, size, and strength of that relationship can be misestimated if unobserved variables that are related to both the independent and dependent variable are omitted. Perhaps the most often cited example is that of the relationship between education and job attainment or income. Both education and job attainment are likely related to individual aptitude or ability, which is difficult to observe or measure in any meaningful way. In a cross-sectional analysis, there isn't a way to disentangle the two to determine the independent effect of additional education by controlling for individual aptitude or ability. A longitudinal analysis, though, which would track the same individual for multiple years, can hold the individual's aptitude as a constant, since it is within that individual, and then observe only the effect of additional years of education on that individual's job attainment or income. In other words, longitudinal analysis allows analysts to control for unobserved factors that may be specific to the individual subject in order to evaluate the independent variable of interest. In political science, the unobserved factors of a specific subject can often be of even greater concern. For example, a nation's unique history, culture, composition, and geography will likely have an important effect on many dependent variables of interest that would be extremely difficult to incorporate into a formal model.

To summarize, there are several advantages of analyzing longitudinal data. It allows for the observation of changes and patterns in changes over time, including within-individual change and interindividual differences in change (Singer & Willett, 2003). It offers the ability to control for heterogeneity created by omitted variables that is typical in political science and economic data by dealing with intersubject variation, and it allows for much greater flexibility in modeling differences in behavior across individuals. Although longitudinal analysis can be a powerful tool for answering a multitude of questions within the discipline, it comes with heavy costs. Collecting the quantity and quality of data that is needed is time-consuming and often expensive. For future waves of data to be collected, the researcher must have a plan and make provisions for the cost of future observations. (Retrospective studies obviously wouldn't face this particular challenge.) Another particular challenge of longitudinal methods is the loss of subjects. Subjects on a panel, particularly individuals who are asked to provide repeated measurements or take repeated surveys, will often drop out for a variety of reasons. Subjects may move out of the area of interest, die, simply stop participating out of boredom or disinterest, or otherwise leave the sample. This attrition can create missing data, which can have significant consequences for analyzing and interpreting the data. This is important to keep in mind when analysts are determining the frequency and length of surveys that will be taken. In some cases, it may be worth sacrificing the additional data that could come from longer surveys or more frequent measurements in order to prevent attrition among participants. In some cases, researchers may also have to worry about the possible

effects of interviewing on the respondent if it will possibly have an impact on independent or dependent variables of interest (Goldfarb, 1960). It is not unreasonable to think that an individual's level of interest in politics, political behavior or activism, and attitudes could all possibly be impacted by the experience of participating in several interviews regarding these subjects.

It is often the case that combining cross-sectional analysis with analysis of available longitudinal data can strengthen the confidence of conclusions in an efficient manner. In their study of the relationship between cumulative voting systems and voter turnout, Bowler, Brockington, and Donovan (2001) combine a cross-sectional and a longitudinal approach. For the cross-sectional analysis, they compare locations with cumulative voting systems with like locations using plurality elections to examine differences in voter turnout. Although they attempt to control as much as possible for differences between communities, their analysis is strengthened by using a longitudinal analysis design to evaluate data for turnout before and after cities changed from plurality systems to cumulative voting systems; they collect data from localities with cumulative voting systems for all of the elections since the conversion and the last three elections prior to transitioning. Data is collected at the city level for fixed time increments since they correlate with election periods. In this case, combining the methods helped to strengthen confidence in the outcomes. Goldfarb (1960) also suggests that it can be useful to conduct a cross-sectional analysis prior to embarking on longitudinal data collections because conducting the less expensive cross-sectional study first can help researchers to identify and develop important concepts and measures earlier on.

Historical Developments

As is the case with many statistical methods, including classic linear regression models, longitudinal analysis can trace its origins back to the field of astronomy (Fitzmaurice & Molenberghs, 2008). Within the last 40 years, methods for longitudinal analysis have advanced substantially, in part because of increased governmental funding for large-scale longitudinal studies and increased computing power and software sophistication. As might be expected, the earliest approaches to analyzing change through longitudinal data were grounded in analysis of variance (ANOVA) methods. In the late 1970s and early 1980s, significant works in the life sciences (Laird & Ware, 1982) and the humanities (Goldstein, 1979) demonstrated the usefulness of ANOVA methods for analyzing longitudinal data and brought increased attention to the development of the methods. As Fitzmaurice and Molenberghs (2008) note, "While ANOVA methods can provide a reasonable basis for a longitudinal analysis in cases where the study design is very simple, they have many shortcomings that have limited their usefulness in applications" (p. 7). To provide

valid estimates, these models required adherence to a number of strict assumptions, including the assumption that variance remains constant over time. The models were designed to apply to experimental research designs where the number of observation points was fixed and common to all individuals. These restrictions were generally untenable with real-world data and models, particularly within political science. As Fitzmaurice and Molenberghs explain,

It was these features of longitudinal data that provided the impetus for statisticians to develop far more versatile techniques that can handle the commonly encountered problems of data that are unbalanced and incomplete, mistimed measurements, time varying and time invariant covariates, and responses that are discrete rather than continuous. (p. 7)

Current Approaches

Linear mixed-effects models are the most common current method for analyzing longitudinal data with a continuous dependent variable. Investigators in the biomedical sciences, education, psychology, and social sciences have contributed to the rapid development of algorithms and extensions of these types of models. As Fitzmaurice, Laird, and Ware (2004) note,

In the early 1980's, Laird and Ware proposed the use of the EM algorithm to fit a class of linear mixed effects models appropriate for the analysis of repeated measurements (Laird & Ware, 1982); Jennrich and Schluster (1986) proposed a variety of alternative algorithms, including Fish scoring and Newton Raphson algorithms. Later in the decade Liang and Zegar introduced the generalized estimating equations in the biostatistical literature and proposed a family of generalized linear models for fitting repeated observations of binary and count data (Liang & Zegar, 1986; Zegar & Liang, 1986). (p. 2)

Today, these models can handle issues of unbalanced data; variables that do and do not vary with time; continuous, binary, and categorical dependent variables; relatively long

or short panels; and a variety of other complications that are beyond the capacity of the earliest models. For an extensive and more thorough review of the evolution and history of these models, see Fitzmaurice and Molenberghs (2008).

Data Considerations

To be useful, longitudinal data must be carefully collected and organized. In cross-sectional analysis, individual data points are generally notated as x_i , with a single subscript indicating the individual. Longitudinal data require that each data point have two identifiers, which can be notated as x_{it} , with an additional subscript t denoting the wave of measurement in time. The error term also would have two dimensions, one for the individual and one for the time period. There are two different arrangements for organizing longitudinal data.

A person level data set has as many records as there are people in the sample. As you collect additional waves, the file gains new variables, not new cases. A person period data set has many more records—one for each person period combination. As you collect additional waves of data, the file gains new records, but no new variables. (Singer & Willett, 2003, p. 17)

Most of the commonly used statistical software packages can convert one form of data into the other with relative ease. Singer and Willett (2003) recommend, though, that when it comes to interpreting and analyzing data, the person-period data set is better for analysis of change over time. The STATA software package also requires that the data be in the person-period format to be recognized as panel data.

An example of the organization of a typical longitudinal data set in person-period format is included below. In this example, states are the individual subjects, and years are the measure of time. There are three independent variables in the example.

<i>State</i>	<i>Year</i>	<i>Dependent Variable</i>	<i>Independent Variable 1</i>	<i>Independent Variable 2</i>	<i>Independent Variable 3</i>
Alabama	2000	$X_{1,1}$	$Y1_{1,1}$	$Y2_{1,1}$	$Y3_{1,1}$
Alabama	2001	$X_{1,2}$	$Y1_{1,2}$	$Y2_{1,2}$	$Y3_{1,2}$
Alabama	2002	$X_{1,3}$	$Y1_{1,3}$	$Y2_{1,3}$	$Y3_{1,3}$
Alabama	2008	$X_{1,9}$	$Y1_{1,9}$	$Y2_{1,9}$	$Y3_{1,9}$
Alaska	2001	$X_{2,1}$	$Y1_{2,1}$	$Y2_{2,1}$	$Y3_{2,1}$

State	Year	Dependent Variable	Independent Variable 1	Independent Variable 2	Independent Variable 3
Alaska	2002	$X_{2,2}$	$YI_{2,2}$	$Y2_{2,2}$	$Y3_{2,2}$
Alaska	2008	$X_{2,9}$	$YI_{2,9}$	$Y2_{2,9}$	$Y3_{2,9}$
Wyoming	2001	$X_{50,1}$	$YI_{50,1}$	$Y2_{50,1}$	$Y3_{50,1}$
Wyoming	2008	$X_{50,9}$	$YI_{50,9}$	$Y2_{50,9}$	$Y3_{50,9}$

Longitudinal data can be experimental, as is the case in many biomedical studies, or observational, as is more often the case in political science. The data can be collected prospectively after the research model has been designed or retrospectively by using existing historical data. Subjects can be individuals, families, organizations, states, and even nations, as is often the case in studies of international relations.

Time can be measured in a variety of ways as well. Pharmaceutical studies interested in the short-term effects of a drug may have observations every 15 minutes for a total of 2 hours. Some longitudinal data sets, however, have multiple years between observations. Most often, longitudinal data sets are wide (large number of subjects) but short (few time periods). Longer longitudinal data sets are now becoming more available, but analysis of these data sets may also have to address additional complications from using nonstationary data, a problem that has long been recognized in time-series analysis (Baltagi, 2005). Stationarity simply means that all the errors have the same expectation and a common variance that does not change over time (Fox, 1997).

There are a few additional ideas about the nature of the data that are worth noting. A balanced longitudinal data set is one in which every individual is observed for the same number of times and at the same intervals. This can often be the case in experimental designs where the entire groups are measured, exposed to a particular treatment—perhaps a political campaign ad—and then measured a few subsequent times. If the time between intervals and the number of individuals is small, it may be relatively easy to ensure that they are all measured at the same intervals. In observational studies, which are more frequent in political science, unbalanced panels are much more common. The panel may be considered unbalanced because not all of the individuals are observed for the same number of times, there are missing data points for particular subjects during certain years, or the subjects are measured the same number of times but at differing intervals.

Although there are statistical methods for modeling and handling missing data, it is important that the analyst understand the nature of the missing data. Missing data is inherently a loss of information and can result in a loss of efficiency or a drop in precision. Perhaps more important, though, under certain circumstances, “the missing data can introduce bias and thereby lead to misleading inferences about changes in the mean response” (Fitzmaurice et al., 2004, p. 376). This is the most serious of the potential problems created by missing data. To know whether bias is likely to be introduced, it is important to understand the reason that the data is missing. If the dropped data points are randomly distributed and have occurred from random attrition among subjects, then it is likely not a concern; however, if those that have dropped out of the study are somehow systematically related to the outcome of interest, then bias is likely. For example, a researcher interested in civic participation may use a longitudinal survey study to evaluate commitment to or interest in civic participation in a sample of individuals. It is possible that those that drop out of the study may do so because they fail to recognize the value of studying civic participation and are likely less interested in civic participation in general. Missing data points that are somehow systematically related to variables of interest require greater care and should be addressed in the study design, or the systematic error should be quantified by the researcher.

Models for Longitudinal Analysis

Depending on the specific nature of the data being used, the research question being addressed, and often the discipline of the analysts, longitudinal analysis can be approached through one of several methodological choices. In some instances, longitudinal analysis is treated as a special case within time series. Bayesian models are still evolving in their ability to be applied to longitudinal data as well. Longitudinal data can also be

treated as a special type of clustered or hierarchical data. Clustered data methods, such as robust regression, are typically used when individuals within the study are likely to be systematically correlated in meaningful clusters. Classic examples within the field of education are classrooms, schools, and districts. Students in the same class are likely to be related on the dependent variable in ways that are not captured by the independent variables in the model. In the case of longitudinal analysis, the clusters are made up of all of the observations for any one individual. The observations are likely to be more closely related to all other observations from that subject than to others.

Since they are perhaps the most commonly employed models within the political science literature, two types of linear regression models deserve particular attention: fixed effects and random effects. There are two different ways of modeling time and individual effects, and the choice between the two models depends on the nature of the data and the theories driving the model. The fixed effects model, also known as the least squares dummy variable (LSDV) model, generates a dummy variable for each of the subject observations across cross-sectional units to control for their individuality (again, think culture, unique history, or individual characteristics that are not incorporated into the model). This model arises from the assumption that the omitted effects are correlated with the included variables (Greene, 2008). Creating the dummy variables controls for this effect and ensures unbiased estimates of the relationship between the independent and dependent variables, but it does so at a heavy cost to efficiency. Each dummy variable results in a loss of a degree of freedom. When the number of subjects is very large, this loss of efficiency is substantial. Essentially, the fixed effects model uses dummy variables to throw out all of the between-subject variance, which is often essential to making causal arguments that need to be purged of third-variable (subject uniqueness) bias. This process also means that the fixed effects model cannot produce estimates for time-invariant variables, such as gender. If the variable does not change over time for the individual subject, or even if it changes very little, the fixed effects model will fail to provide precise coefficient estimates.

Because the random effects model is both more efficient and able to estimate time invariant variables, it is used most often. However, the random effects model can produce bias in the error terms, in which essentially a measure of all the latent or unmodeled omitted variables one tries to capture without modeling are systematically related to the dependent variable or the relationship between the independent and dependent variables. In other words, where states are the subject, a researcher needs to know if there is something about South Dakotans that he or she is not including in the model that will be correlated across all of the error components for South Dakota and

that is also expected to be correlated with the independent variables in the model. One should address this question foremost with theory and an understanding of the data, but it can also be tested for empirically with the Hausman test.

The Hausman test considers the null hypothesis that the difference in the coefficients is not systematic. As Yaffee (2005) notes,

The test for this correlation is a comparison of the covariance matrix of the regressors in the LSDV (Least Squares Dummy Variable) model with those in the random effects model. The null hypothesis is that there is no correlation. If there is no statistically significant difference between the covariance matrices of the two models, then the correlations of the random effects with the regressors are statistically insignificant. (p. 8)

If the null hypothesis is rejected, then the random effects model may be biased and is not appropriate (Gujarati, 2003), but if the null hypothesis is not rejected, the random effects model is more efficient and preferred because it allows for time invariance in the independent variables.

Extensions of Longitudinal Analysis

The models discussed so far have primarily focused on longitudinal analysis with a continuous dependent variable. Extensions of the longitudinal analysis model have been developed that allow for analysis of dichotomous and categorical data as well. The last 30 years have witnessed a substantial advancement of the methods for analyzing discrete longitudinal data in part because of the ability to readily harness high-speed computing resources (Fitzmaurice & Molenberghs, 2008). Over the years, many extensions of longitudinal analysis have been developed. Some are extensions of generalized linear models (GLMs). GLMs are a “unified class of models for regression analysis of independent observations of a discrete or continuous response” (Fitzmaurice & Molenberghs, p. 9). Likelihood-based approaches have also been explored but have found the extension to longitudinal analysis of discrete data somewhat more challenging, but alternatives have made progress. Greene (2008, chap. 23) offers a nice overview of explanations and applications for discrete choice models with respect to panel data, and for a more extensive review, Molenberghs and Verbeke (2005) is a good resource.

Event history analysis, also called *survival analysis* or *duration analysis*, is a special type of panel data analysis used primarily within the study of international relations and comparative politics where organizations, states, or governments can be the unit of analysis. In these models, rather than observations being scheduled at specified intervals, measurements are taken at each stage of a sequence of events. As Box-Steffensmeier (1997) notes,

Time plays a key role in politics and a class of econometric models, known collectively as event history analysis, can provide researchers leverage on the issue of the timing of political change. Event history analysis allows researchers to answer a more extensive set of questions than conventional analyses by using information on the number, timing, and sequence of changes in the dependent variable. (p. 1414)

Many of the earliest time-to-event data analyses were conducted in the biomedical or health sciences fields and often focused on morbidity. The publication of Cox's seminal paper on proportional hazards models in 1972 contributed significantly to the method's further development (Fitzmaurice et al., 2004). Today, event history analysis is used in a variety of fields to study the occurrence of distinct events such as wars, loan defaults, or significant political changes. In his article, Gasiorowski (1995) uses event history analysis to examine the relationship between economic crises and political regime change. In this panel data set, observations for a created dichotomous dependent variable that indicates whether a regime change occurred during a given year are taken for 97 countries. For more information on the specifications of these event history models, see Box-Steffensmeier and Jones (2004) and Singer and Willett (2003).

Statistical Programs

The increasing popularity of longitudinal analysis with multiple disciplines has led to the development and expansion of statistical options in almost all of the major statistical software programs. STATA, SAS, and LIMDEP all have the capacity to run fixed and random effects models, can handle the additional complexity of unbalanced panels, have options for one-way or two-way random and fixed effects models, and can correct for autocorrelation. LIMDEP and STATA also have procedures for limited dependent panel data analysis including negative binomial, logit, and probit. Both programs also provide models for analyzing panel stochastic frontier data (Yaffee, 2005). With the latest versions, Releases 10 and 11, STATA has published an entire manual devoted to the analysis of longitudinal and panel data. STATA provides options for an impressive variety of longitudinal models including the Arellano-Bond linear dynamic panel data model, linear panel data models for fixed effects, random effects and population averaged models, models for logit and probit analysis of panel data, stochastic frontier models, and Poisson models for panel data. The program also includes a broad range of hypotheses tests, diagnostics, graphing capabilities, and postestimation techniques. In 2008, Croissant and Millo introduced an add-on package called *plm*, which makes data management and estimation of linear panel data models fairly straightforward in R, an open-source software platform for statistical analysis. (For a detailed description and information on downloading the

software, see Croissant & Millo, 2008.) There are also specialized software packages that were designed specifically for fitting multilevel models for change to data including HLM (Raudenbush, Bryk, Cheong, & Congdon, 2001), which is the program of choice for examples in Singer and Willett's (2003) instructional text, *Applied Longitudinal Data Analysis*. With so many options, the choice is in large part user preference. There is some evidence that the different programs produce the same or very similar results (Singer & Willett), so analysts may choose between them based on familiarity with the overall program, ease of use, look of visual outputs, or software availability.

Future Directions

The same characteristics that make longitudinal data analysis so rich and interesting also create analytical complexities for statistical models. For example, the random effects model, which allows for time invariant variables and greater efficiency, is often the most applicable for the types of data and research questions analyzed in political science research, but it is also the most susceptible to bias. These models continue to evolve as methodologists continue to expand and refine these methods to deal with greater complexity and to incorporate these solutions into statistical software, which is also ever increasing in capacity, so that they are available for scholars and practitioners in the field. "Thus, the outlook is bright that modern methods for longitudinal analysis will be applied more widely and across a broader spectrum of disciplines" (Fitzmaurice & Molenberghs, 2008, p. 3).

An additional factor contributing to the increasing use of longitudinal methods is the increasingly large amount of longitudinal data available for analysis. Large-scale longitudinal data sets of interest used to be relatively hard to come by, but now panel data sets are available on a multitude of topics, and more continue to be developed. One prominent example of longitudinal data that can be applicable to research questions within the field is the Panel Study of Income Dynamics (PSID). As the study authors note, "The PSID is a nationally representative longitudinal study of nearly 9,000 U.S. families. Following the same families and individuals since 1968, the PSID collects data on economic, health, and social behavior" (Institute for Social Research, n.d., para. 1). Impressive in both its size and its quality of information, the National Longitudinal Study of Youth (NLSY) is "a nationally representative sample of 12,686 young men and women who were 14–22 years old when they were first surveyed in 1979" (Bureau of Labor Statistics, n.d., para. 1). The study continued to interview the same individuals annually through 1994, providing 15 years of annual data, and currently interviews them every other year on employment, marital status, economic status, education, drug use, and health. In 2007 and 2008, the American National Election Studies (ANES)

incorporated for the first time a 2-year panel study, which interviewed individuals in late 2007—before the primaries, in the months running up to Election Day, in November 2008, and in May 2009. As ANES (n.d.) authors note, “The panel will allow scholars to study citizen politics in new ways and will illuminate how election year politics affect judgments of the new administration in the formative months of its term” (para. 4). The Correlates of War Project offers numerous longitudinal data sets that are frequently used within the study of international relations. This is just a short list of a rapidly growing collection of longitudinal data sets that are available for analysis within the discipline. However, with greater availability of data and easier to use programs, the likelihood for misuse of these analytical models will increase as well. Wilson and Butler (2007) identify 195 studies published in political science journals that employ linear panel data methods to evaluate whether the authors consider unit heterogeneity and dynamic specifications and found only 5% of the studies met their limited criteria. They find a general “lack of attention to specification issues and a failure to adequately consider well-known models found in the literature” (p. 110). Hence, it is critical that as these models become more popular and are used with greater frequency, more time is spent educating scholars within the discipline on their proper use and interpretations.

Conclusion

An increasing number of longitudinal studies and panel data collection efforts as well as the increased ability to track individual subjects over time have provided the impetus for analysts to harness the advantages of greater computing power in order to capitalize on the benefits offered by analyzing longitudinal data. By combining a dynamic element with the richness of cross-sectional observations, panel studies give “more informative data, more variability, less collinearity among variables, more degrees of freedom and more efficiency” (Gujarati, 2003). Developments in the analysis and collection of longitudinal data in the last 50 years have allowed political science scholars to address questions that were previously difficult to answer with much confidence or certainty. Greater computing power and the advances of statistical software along with the exploding growth of longitudinal data sets have made longitudinal data analysis one of the fastest growing methods in the field. Researchers now have a variety of tools at their disposal for dealing with a whole spectrum of complex questions and data. The popularity of the methods will likely continue to increase the availability and precision of these models. Although this offers much promise, it is also important that analysts take time to understand the differences, assumptions, and needs of the various models. This chapter has provided only a brief glimpse of the

distinctions between available options; however, the references and further readings list offers several volumes that provide the technical, theoretical, and practical knowledge needed for the application and further development of these powerful tools.

References and Further Readings

- American National Election Studies. (n.d.). *About ANES*. Ann Arbor, MI: Center for Political Studies. Retrieved May 15, 2009, from <http://www.electionstudies.org/overview/overview.htm>
- Arellano, M. (2002). *Panel data econometrics*. Oxford, UK: Oxford University Press.
- Baltagi, B. (2005). *Econometric analysis of panel data* (3rd ed.). New York: Wiley.
- Bowler, S., Brockington, D., & Donovan, T. (2001). Election systems and voter turnout: Experiments in the United States. *Journal of Politics*, 63(3), 902–915.
- Box Steffensmeier, J. M. (1997). Time is of the essence: Event history models in political science. *American Journal of Political Science*, 41(4), 1414–1461.
- Box Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling: A guide for social scientists*. New York: Cambridge University Press.
- Bureau of Labor Statistics. (n.d.). *National longitudinal surveys: NLSY79*. Washington, DC: Author. Retrieved May 15, 2009, from <http://www.bls.gov/nls/nlsy79.htm>
- Croissant, Y., & Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2), 1–43.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: John Wiley & Sons.
- Fitzmaurice, G. M., & Molenberghs, G. (2008). Advances in longitudinal data analysis: An historical perspective. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis handbooks of modern statistical methods* (pp. 3–27). Boca Raton, FL: CRC Press.
- Fox, J. (1997). *Applied regression analysis, linear models, and related methods*. Thousand Oaks, CA: Sage.
- Gasiorowski, M. J. (1995). Economic crisis and political regime change: An event history analysis. *American Political Science Review*, 89(4), 882–897.
- Goldfarb, N. (1960). *Longitudinal statistical analysis: The method of repeated observations of a fixed sample*. Glencoe, IL: Free Press.
- Goldstein, H. (1979). *The design and analysis of longitudinal studies*. London: Academic Press.
- Greene, W. H. (2008). *Econometric analysis* (6th ed.). Saddle River, NJ: Pearson Education.
- Gujarati, D. N. (2003). *Basic econometrics* (4th ed.). New York: McGraw Hill.
- Institute for Social Research. (n.d.). *Panel study of income dynamics*. Ann Arbor: University of Michigan. Retrieved May 13, 2009, from <http://psidonline.isr.umich.edu>
- Jennings, M. K. (2002). Generation units and the student protest movement in the United States: An intra- and intergenerational analysis. *Political Psychology*, 23(2), 303–324.

- Jennings, M. K., & Niemi, R. G. (1968, July). The transmission of political values from parent to child. *American Political Science Review*, 62(1), 169–184.
- Jennings, M. K., & Niemi, R. G. (1978, July). The persistence of political orientations: An over time analysis of two generations. *British Journal of Political Science*, 8(3), 333–363.
- Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics*, 38(4), 963–974.
- Liang, K. Y., Zeger, S. L., & Qaqish, B. (1992). Multivariate regression analyses for categorical data (with discussion). *Journal of the Royal Statistical Society, Series B (Methodological)*, 54(1), 3–40.
- Lubotsky, D. (2007). Chutes or ladders? A longitudinal analysis of immigrant earnings. *Journal of Political Economy*, 115(5), 820–867.
- Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data*. New York: Springer.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y., & Congdon, R. (2001). HLM [Computer software]. Lincolnwood, IL: Scientific Software International. Available from <http://www.ssicentral.com>
- Singer, J. B., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Wilson, S. E., & Butler, D. M. (2007). A lot more to do: The sensitivity of time series cross section analyses to simple alternative specifications. *Political Analysis*, 15(2), 101–123.
- Woolridge, J. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.
- Yaffee, R. (2005). A primer for panel data analysis. *Connect*. New York: New York University. Retrieved May 29, 2009, from <http://www.nyu.edu/its/pubs/connect/fa1103/yaffeeprimer.html>

60

QUALITATIVE VERSUS QUANTITATIVE RESEARCH

ALICIA JENCIK

University of New Orleans

For decades, there has been a raging debate among scholars regarding the differences between and advantages of qualitative and quantitative methods. In fact, this has probably been one of the largest and longest methodological debates in all of social science research. Perhaps it can be briefly summarized by the following two famous and opposing quotations: Donald Campbell says, “All research ultimately has a qualitative grounding”; and Fred Kerlinger says, “There’s no such thing as qualitative data. Everything is either 1 or 0” (in Miles & Huberman, 1994, p. 40). Although it is not necessarily critical to determine which—if either—of these approaches can be described as the better one, it is imperative to have a thorough understanding of these methods in order to be able to conduct sound political science research. After all, for a study to be of value to scholars and other individuals interested in the topic, it is necessary for one to choose the correct research approach, ask suitable questions, use appropriate research methods and statistical analyses, correctly deduce or induce inferences, and have suitable general goals driving the research.

The questions under consideration and the answers obtained by any particular study will depend on whether the study uses quantitative or qualitative approaches. The purpose of this chapter is to differentiate between these two types of research. First, the literature available on this topic is briefly summarized, focusing specifically on how qualitative and quantitative research is defined, as well as the

different assumptions on which these types of research are based. Next, a summary of the similarities and differences in each stage of the research process is provided. Then, the different methods that these two types of approaches use are discussed. Next, since this is a book examining political science in the 21st century, current and future research directions are examined. In particular, the use of what are called *mixed methods* approaches is discussed. The chapter ends with a brief summary and conclusion of the information that has been presented. Finally, suggested books and articles for further reading are provided, including some material for individuals interested in conducting advanced statistical studies, which are beyond the scope of this chapter.

Quantitative and Qualitative Research

The following section introduces the definitions and assumptions of quantitative and qualitative research. First, however, it is worth briefly discussing two types of political analysis in order to understand the origins of quantitative and qualitative methods. Political scientists distinguish between empirical analysis—obtaining and dealing with knowledge and information—and normative analysis—determining how to use that knowledge. Normative analysis relies on the development of subjective goals and values to apply what has been learned to reality. Empirical analysis, however, focuses on using common terms to explain and

describe political reality and can be either quantitative or qualitative in nature. If something is empirical, it is verifiable through observations or experiments. Empirical analysis is the focus of this chapter.

Definition of Quantitative Research

As a first step, it is necessary to define these two methods of research and examine their goals. Quantitative research can be defined as a process of inquiry examining an identified problem that is based on testing a theory measured by numbers and analyzed with statistical techniques. Thus, quantitative research involves the analysis of numerical data. A more technical definition is provided by Brady and Collier (2004), who define mainstream quantitative methods as “an approach to methodology strongly oriented toward regression analysis, econometric refinements on regression, and the search for statistical alternatives to regression models in contexts where specific regression assumptions are not met” (p. 294). The econometric refinements and statistical alternatives referred to by the authors are beyond the scope of this chapter but include logit and probit models, time-series analysis, and a variety of techniques to circumvent problems that can occur in regression analysis, such as heteroskedasticity and autocorrelation. Essentially, quantitative methods have played a major role in improving on commonly used research tools within the structure of regression models that are frequently used in the field of political science.

Assumptions of Quantitative Research

The goal of quantitative research is to examine particular instances or aspects of phenomena to determine if predictive generalizations of a theory hold true or to test causal hypotheses. As a result, there are several key assumptions underlying quantitative research methods, which are briefly outlined here. These include the following:

- Reality can be studied objectively.
- Research must remain independent of the researcher through the use of experiments, questionnaires, machines, or inventories.
- Research is value free, and the researcher does not become a part of or interfere with the research.
- Theories and hypotheses are tested in a cause-effect order with research based primarily on deductive forms of logic identified a priori by the researcher.
- The purpose of research is to develop generalizations that contribute to theory and allow the researcher to predict, explain, and understand a particular phenomenon.

Definition of Qualitative Research

Qualitative research can be defined as a process of inquiry that builds a complex and holistic picture of a

particular phenomenon of interest by using a natural setting. Thus, qualitative research involves the analysis of words, pictures, videos, or objects in the context in which they occur.

Assumptions of Qualitative Research

The goal of qualitative research is to understand social issues from multiple perspectives to have a comprehensive understanding of a particular event, person, or group. As with quantitative research, there are several key assumptions underlying qualitative research methods:

- Reality is socially constructed, and there are multiple realities.
- The researcher interacts and often works closely with the individuals or groups under study and serves as the primary instrument for data collection and analysis.
- The research is value laden, and the researchers become a part of the research, attempting to understand the lives and experiences of the people they study.
- Research is context bound and based on inductive forms of logic that emerge as a study progresses.
- The purpose of research is to find theories that help explain a particular phenomenon.

Comparing and Contrasting Quantitative and Qualitative Research Methods

The following section examines how quantitative and qualitative methods are similar to and different from each other throughout the research process, beginning with the creation of a research question and up to the reporting of the results. Although examining quantitative and qualitative methods as two separate categories is necessary for the sake of clarification throughout this section, it is important to realize that these two methods are not mutually exclusive, a topic that will be discussed in more detail shortly. As Manheim, Rich, Willnat, and Brians (2007) note, when examining the differences between quantitative and qualitative methods, “The distinctions discussed are generally more matters of degree than absolutes. The two types of methods often require only different *forms* of work, but are working toward similar objectives” (p. 323). This is important to keep in mind while reading this chapter.

The Research Question

The first step in conducting sound political science research is selecting a research question. An appropriate research question should fulfill either a scientific need or a societal need by helping to provide an answer to an important problem. Both quantitative and qualitative forms of research begin by creating a research question that is

intended to produce knowledge of the empirical world. In terms of the research questions, the main difference between quantitative and qualitative methods typically exists in the type of questions that are being posed.

Theorizing

A theory is a potential explanation for events and is composed of a set of logically related propositions and assumptions. Theorizing is the actual process of stating these conceptual explanations for events that take place in the real world by proclaiming relationships among the concepts. Theories are created to help people understand phenomena. There are several characteristics that make a theory particularly useful in explaining observations. Theories should be (a) testable, (b) logically sound, (c) communicable, (d) general, and (e) parsimonious.

Theorizing is a critical phase of the research process for quantitative and qualitative researchers. However, quantitative researchers are more likely than qualitative researchers to focus on testing performed theories. Quantitative researchers base their studies on a theory that relates to their subject in an attempt to develop generalizations that contribute to theory. Thus, in quantitative research, theorizing occurs prior to the collection of data. Qualitative researchers, on the other hand, are more likely than quantitative researchers to elaborate on theories while making observations of a particular phenomenon. Many qualitative researchers argue that, as a result of this, their theories are far more grounded in reality than are those of quantitative researchers. However, quantitative researchers argue that the formulation of theory during the observation-making process can easily lead to the creation of a theory designed around those specific observations. As a result, these theories would be polluted and not testable. Furthermore, if a theory is based on observation of one particular group, the usefulness of the theory is quite limited.

Research Design

Simply defined, a research design is the plan of a study. It organizes observations in a manner that establishes a logical basis for causal inference. Essentially, the research design can be viewed as the blueprint for a study. There are three main types of research designs in political science: exploratory, descriptive, and explanatory. Exploratory research attempts to discover which factors should be included when theorizing about and researching a particular subject. Descriptive research attempts to measure some aspect of reality for its own sake and not for the purpose of developing or testing some theory. Explanatory research uses observations of reality to test hypotheses and help develop an understanding of patterns of behavior in the context of a specific theory.

Regardless of the purpose of a study, every research design should have the same basic elements, which are outlined by Manheim et al. (2007): (a) a statement outlining

the purpose of the research; (b) a review of the theory and any hypotheses that are going to be tested, if applicable; (c) a statement explaining the variables that will be used; (d) an explanation of the operationalization and measurement of the variables; (e) a statement of how observations will be organized, as well as conducted; and (f) a discussion of how the data that are collected will be analyzed.

Although both quantitative and qualitative researchers produce research designs for their studies, quantitative researchers are much more likely than their counterparts to base their designs on the logic of experiments. For instance, quantitative researchers often emphasize control groups, pretests, and other elements that provide them with the opportunity to hold some factor(s) constant in their attempt to make causal inferences. Qualitative research designs, on the other hand, typically focus more on who or what is being observed, where the observation will take place, how observations will be conducted, and how the data will be recorded. For qualitative researchers, more emphasis is placed on viewing people and events as they naturally occur, while for quantitative researchers there is a greater focus on establishing cause-and-effect relationships.

Sampling

A sample is a small group of cases drawn from and used to represent a larger population under consideration. A representative sample is a sample in which each major attribute of the larger population occurs in approximately the same proportion or frequency as in the larger population. "In other words, a truly representative sample is a microcosm—a smaller, but accurate model—of the larger population from which it is taken" (Manheim et al., 2007, p. 119). When a sample is representative, the conclusions drawn from it are generalizable to the entire population.

In quantitative studies, sampling is based on the logic of probability to produce statistical representativeness. Additionally, in quantitative research, sampling is done before the data are collected. Qualitative researchers, on the other hand, usually create their sample once their study is already in progress. After observing, learning about, and gaining understanding from an initial case, qualitative researchers are then able to determine what they will observe next. Additionally, whereas generalizability is a chief concern for quantitative researchers, this is not the case for qualitative researchers, who are far more concerned with finding the specific information that they are looking for from their sample. Since this method is very time-consuming, qualitative findings are often based on fewer cases than quantitative findings.

Data Collection

Data are observations or information about reality that represent attributes of variables and result from the research process. Although data collection is an integral

part of both types of research methods, data are composed of words in qualitative research and numbers in quantitative research, which results in a data collection process that differs significantly for quantitative and qualitative research. Furthermore, the data collection process is different: Although quantitative researchers have the ability to administer a previously prepared questionnaire or watch an experiment unfold behind blind glass, qualitative researchers are engaged—sometimes for long periods of time—with the people or groups under study.

Data Analysis

As can likely be seen by now, quantitative researchers frequently have a detailed plan of action that is thought out prior to the beginning of a study's taking place. Qualitative researchers, on the other hand, tend to take a more fluid approach to their studies. This holds true for the analysis of data, as well. Whereas in quantitative studies, the data analysis methods are planned out in advance and then occur after the data are collected, data analysis typically takes place at the same time as data collection in qualitative studies. To make appropriate future observations, analyses must often begin after studying one to several initial cases. As a result, quantitative researchers are not usually afforded the opportunity to modify their methods of data collection during a project, while qualitative researchers can do so at any point in a project after conducting the initial data analysis.

Additionally, although qualitative data are more subjective and sometimes difficult to interpret, quantitative data are easily coded into numerical formats. As a result, it is much easier to enter quantitative data into computer programs, such as Excel and SPSS, than it is to enter qualitative data. Furthermore, there are a number of programs that analyze the statistical data, such as SPSS and Stata. Although programs do exist for the interpretation of qualitative data, they are not used nearly as extensively as those used for quantitative data analysis.

Finally, whereas quantitative researchers have a variety of means to test the statistical significance and validity of the data that they are analyzing, this is not the case for qualitative researchers. Instead, qualitative researchers must do their best to present a clear, accurate, and convincing analysis of their data. As a result, a topic of much debate between quantitative and qualitative researchers is the validity and reliability of findings produced in studies. Validity is the extent to which measures correspond to the concepts they are intended to reflect. Reliability is the consistency with which a measuring instrument allows assignment of values to cases when repeated over time. Although a measure can be reliable without being valid, it cannot be valid without being reliable.

Additionally, since one of the main points of conducting quantitative research is to study causal relationships, part of the process involves manipulating various factors that could potentially influence a phenomenon of interest

while at the same time controlling for other variables that could affect the outcome. For instance, if a researcher were examining if gender played a role in whether a person received a job, it would be important to control for other variables, such as education or previous work experience, since these factors may also determine why an individual would receive an employment offer. In quantitative analysis, empirical relationships and associations are typically examined by using general linear models, nonlinear models, or factor analysis to understand important information about the relationship between variables, such as the direction of a relationship. However, despite the results that may be produced by these models, it is important to note that a major tenet of quantitative research is that correlation does not imply causation. In other words, a spurious relationship is always a possible result of the data analysis.

Reporting of Results

When presenting the results of a study, qualitative researchers often have an arduous task in front of them. Since their reports typically rely on the interpretation of observations, it is necessary for them to be very careful in the selection of what stories, quotations, pictures, and so on, they will share in order to avoid bias. The reports produced by quantitative researchers tend to be more straightforward since they rely mostly on the interpretation of statistics. But here, too, it is important to make sure that bias was avoided in the sample and that appropriate data analysis methods were used in order to avoid bias in quantitative analysis.

Summary

To sum up, there are a lot of similarities among quantitative and qualitative research methods. Irrespective of which method is used, it is still necessary to create an appropriate research question, understand the theory behind what will be observed, create a research design, collect and analyze data, and create a report of the results. However, there are several key differences between quantitative and qualitative research methods. These methods differ in (a) the types of questions that they pose, (b) their analytical objectives, (c) the amount of flexibility allowed in the research design, (d) the data collection instruments that are used, and (e) the type of data that are ultimately produced. According to Mack, Woodsong, MacQueen, Guest, and Namey (2005), the fifth difference is the biggest. The authors argue that quantitative methods are generally inflexible since categories are typically closed-ended or fixed, while qualitative methods are more flexible, with a large amount of spontaneity and adaptation occurring during interaction with other people, especially in the form of open-ended questions.

To decide which research approach should be used, several things should be taken into account, including the problem

of interest, the resources available, the skills and training of the researcher(s), and the audience for the research. Since there are considerable differences in the assumptions that underlie these two research approaches, as well as the collection and analysis of data, these considerations are important. The following sections provide a more detailed examination of the various types of quantitative and qualitative research methods, as well as the limitations of these methods in general.

Quantitative Methods

Quantitative methods are essentially a variety of research techniques that are used to gather quantitative data. There are a variety of different types of quantitative methods, which are briefly outlined in this section: experiments, quasi experiments, content analysis, and surveys. First, in experiments, participants are randomly assigned to experimental conditions, as well as experimental controls. The individuals who are assigned to experimental controls are testing the independent variable. The difference between experiments and quasi experiments is the way that subjects are selected. In quasi experiments, participants are assigned to experimental conditions in a nonrandom fashion.

Next, content analysis is a systematic means of counting and assessing information in order to interpret it. For instance, scholars may count the number of times that personal characteristics, such as dress or hairstyle, are mentioned in newspaper articles to determine whether media coverage of male and female candidates differs. Finally, surveys are used to estimate the characteristics of a population based on responses to questionnaires and interviews from a sample of the population. Surveys provide five types of information: (1) facts, (2) opinions, (3) perceptions, (4) attitudes, and (5) behavioral reports. Essentially, questionnaires and surveys can serve as a means for helping scholars understand why people feel or act the way that they do, as well as measure their attitudes and assess their behaviors.

Limitations of Quantitative Methods

There are three key criticisms of quantitative research that are discussed here. First, since quantitative research methods were adopted from the physical sciences, critics argue that all cases are treated as though they are alike. Complex concepts are turned into numbers, and their unique elements are dissipated as a result. Furthermore, people can easily attribute different meanings to something even when they are experiencing the same phenomena. Second, and related to the first criticism, some people argue that quantitative methods are inherently biased. Since they are adopted from the physical sciences, critics argue that quantitative methods fail to take into account the

unique cultural roots and other critical aspects of marginalized groups of people. Thus, according to critics, when it comes to populations that have been politically excluded, the usage of quantitative methods may not be appropriate, according to critics. Third, critics argue that quantitative research methods result in taking individuals out of their natural settings to examine very limited aspects of what a person thinks or believes. To these critics, context is very important, and by taking actions out of context, it is impossible to understand the true meaning of events or responses.

Qualitative Methods

Just as quantitative research methods have a variety of research techniques that are used to gather data, there are also a variety of qualitative methods. This section focuses on several of these: ethnographic studies, phenomenological studies, case studies, focus groups, and intense interviews. First, in ethnographic studies, researchers examine cultural groups in their natural setting. Examples of cultural groups can include students in a dormitory, women in a crisis center, or people from a village in Asia. This type of study can provide rich, detailed information about the individuals in various groups, since it involves first-hand observation.

Second, in phenomenological studies, a small group of people is studied intensively over a long period to understand the life experience of the individuals being studied. Phenomenological studies can involve direct or indirect observation. Additionally, depending on the study, the individuals being observed may or may not know the purpose of the study or what exactly is being observed. Sometimes the researcher relies on building a trusting relationship with the subjects so the subjects act as naturally as possible even though they are being observed. As a result of this closeness, the researcher can often tell when a person is modifying his or her behavior. However, it is not always possible to establish this kind of relationship. As a result, some researchers conceal the purpose of their studies from those being observed to avoid the modifying of behavior by the subject. This process of behavior modification by the respondent is called *reactivity* and can greatly affect the results of a study.

Third, in a case study, a case is studied by a researcher, and detailed information about the entity or phenomenon is recorded. Sometimes information that is found in a case study can lend itself to the content analytical techniques discussed in the previous quantitative research section. Other times, newspapers, books, interviews, or other sources may be used. In content analysis, researchers are looking for specific words, phrases, or general ideas that are relevant to their study. The researchers will then count the instances of these items to learn more about a particular subject. For instance, some political scientists are interested in learning

about gender bias in the media. By examining how often a female versus a male candidate is mentioned in an article or the type of coverage the candidate receives, these scholars are able to draw conclusions about gender bias in the media.

Finally, there are two other ways to collect and analyze qualitative data that are of relevance in this section—focus groups and intense interviewing. Focus groups are in-depth studies composed of small groups of people who have guided discussions. For instance, a focus group may be shown a political advertisement that a political campaign hopes to air on television. After watching the advertisement, members of the group are asked questions, and a discussion is prompted in which they can discuss their feelings about the ad, such as what they liked and did not like, as well as whether they were swayed by the ad and found it to be credible. These responses allow the advertisement's producers to make changes that make the ad more effective.

Intense interviews are similar to survey questionnaires in that the interviewer generally has some thoughts in mind about what the respondent will be asked. However, although survey questions are planned out in their entirety in advance, this is often not the case in intense interviews where the interviewee has the ability to ask follow-up questions or a variety of other questions related to an answer provided by the respondent. Additionally, whereas survey questionnaire responses tend to be closed-ended (a particular response can be chosen from those available), intense interview responses are typically open-ended (no response categories) and can be very detailed. Thus, researchers have more flexibility when conducting an intense interview than they would if they were administering a questionnaire; however, their results are typically not quantifiable.

Limitations of Qualitative Methods

Just as quantitative methods have their detractors, so too do qualitative methods. Some of the biggest criticisms of qualitative methods are outlined in this section. First, some critics argue that qualitative methods focus too much on particular individuals, sometimes at the expense of seeing the bigger picture, and they fail to make their results generalizable to a larger population. Second, critics note that the quality of the results and analysis that are produced are highly dependent on the skill of the researcher. It is necessary for the researcher to have remained unbiased and provide a clear assessment of the subjects under study, or the results are essentially meaningless. Third, it is very time-consuming to conduct qualitative research studies. The amount of time spent conducting interviews and making observations is just the beginning. After these take place, the researchers still have to figure out a way to analyze the vast amounts of information that they have collected to produce results.

Future Directions

As can be seen from the information provided throughout this chapter, there has been a raging decades-long debate as to whether qualitative or quantitative research is better. Many scholars focus on qualitative versus quantitative techniques, automatically framing these methods and approaches in opposition to each other. Although it may appear that qualitative and quantitative data exist in opposition to each other, this is not necessarily the case. As King, Keohane, and Verba (1994) argue, "The two traditions appear quite different; indeed they sometimes seem to be at war. Our view is that these differences are mainly ones of style and specific technique. The same underlying logic provides the framework for each research approach" (p. 3). As a result, research does not typically fit into one particular category or another.

Additionally, King et al. (1994) note that we live in a world that changes rapidly, and to fully understand the changes that occur around us, it is necessary to be able to take into account information that can be quantified, as well as information that cannot. Furthermore, since social science requires comparison, it is important to examine both quantitative differences (such as which phenomena are more or less alike in degree) and qualitative differences (such as which phenomena are more or less alike in kind).

In recent years, scholars have been focusing a lot more on triangulation. Triangulation is essentially the idea that more than one research technique can be used to examine a research question to further verify the findings. Triangulation can help improve confidence about the results produced from a study. Quantitative and qualitative research can frequently be integrated, creating mixed-methods research that can depict a clearer picture of a social science phenomenon than one single method on its own.

Another way that quantitative and qualitative methods can exist together is by coding qualitative data into quantitative data. Just about any type of qualitative data can be assigned meaningful numerical values that can be manipulated to help condense the information and gain a different and more generalizable understanding of the data. One frequently used example is open-ended questions. Although more detailed insight is gained from an open-ended question than a categorical question, open-ended questions can typically be broken down into simple numerical categories allowing for a quantitative analysis of the data.

The Research Network on Gender Politics and the State (RNGS) serves as another good example. The researchers in RNGS had been conducting a cross-national, longitudinal, qualitative research project that explored changes in public policy processes dating back to the 1960s. Starting in 2000, however, the researchers began to code their vast qualitative data into a large quantitative data file. By using quantitative coding, additional

useful information may be garnered, and a new form of data analysis is possible. As can be seen here, sometimes the line between quantitative and qualitative analysis may not be so clear after all.

On the other hand, quantitative data is inherently based on qualitative judgment because it is impossible to interpret numbers without understanding the assumptions underlying the numbers. When a person provides a numerical response to a survey question, for instance, many assumptions and judgments are present. For instance, if a person, when asked, "How satisfied are you with your life?" responds, "Very satisfied" (denoted by a value of 1), a variety of other questions could be asked. What does satisfaction mean to this respondent? Was he or she thinking only of the economic climate? Job? Family? Relationships? How does he or she define satisfaction, and how does this differ from how the next person defines satisfaction? Did the respondent even pay attention to or think about the question, or was he or she just offering quick responses? When and in what context was this question presented? The list goes on. As can be seen from this brief example, what appeared to be a simple numerical piece of information actually involved numerous judgments about the meaning of each response.

Conclusion

Quantitative and qualitative analysis are two general approaches to the analysis of data. Both seek to explain trends but have different means of doing this. Additionally, quantitative and qualitative research methods are each based on a basic set of assumptions. Both forms of research carefully follow each step in the research process, from formulating a research question to reporting the results of the data analysis. However, the order and ways in which this process is completed differ between quantitative and qualitative methods because of the different goals that researchers using these methods have for their studies. Essentially, though, at some level, quantitative and qualitative data are inseparable and do not exist in complete opposition to each other. Thus, it is almost self-defeating to claim that one method is better than the other. There are times when one is more appropriate to use in a given situation than another, but often, they can both be used together, whether at the same time or in different stages. As research progresses through the 21st century, it is highly probable that more scholars will use mixed-methods approaches.

References and Further Readings

- Achen, C. H. (1982). *Interpreting and using regression*. Beverly Hills, CA: Sage.
- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95, 529-546.
- Agresti, A., & Finlay, B. (1997). *Statistical methods for the social sciences* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Berg, B. L. (2003). *Qualitative research methods for the social sciences* (5th ed.). Boston: Allyn & Bacon.
- Box Steffensmeier, J. M., Brady, H. E., & Collier, D. (Eds.). (2008). *The Oxford handbook of political methodology*. New York: Oxford University Press.
- Brady, H., & Collier, D. (Eds.). (2004). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham, MD: Rowman & Littlefield.
- Creswell, J. W. (2009). *Research design: Qualitative and quantitative approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Denzin, N. K., & Lincoln, Y. (2005). *The SAGE handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Fink, A. (2005). *How to conduct surveys: A step by step guide* (3rd ed.). Thousand Oaks, CA: Sage.
- Frankfort Nachmias, C., & Nachmias, D. (1999). *Research methods in the social sciences* (6th ed.). New York: St. Martin's.
- Frey, J. H., & Oishi, S. M. (2004). *How to conduct interviews by telephone and in person*. Thousand Oaks, CA: Sage.
- Gerring, J. (2001). *Social science methodology: A critical framework*. Cambridge, UK: Cambridge University Press.
- Goodin, R. E., & Klingemann, H. D. (Eds.). (1996). *A new handbook of political science*. New York: Oxford University Press.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Towards a conceptual framework for mixed method evaluation designs. *Education Evaluation and Policy Analysis*, 11, 255-274.
- Gujarati, D. N. (2002). *Basic econometrics* (4th ed.). New York: McGraw Hill.
- Hewson, C., Yule, P., Laurent, D., & Vogel, C. (2002). *Internet research methods: A practical guide for the social and behavioural sciences*. Thousand Oaks, CA: Sage.
- Hoover, K., & Donovan, T. (2004). *The elements of scientific thinking* (8th ed.). Belmont, CA: Wadsworth.
- Isaak, A. C. (1975). *Scope and methods of political science: An introduction to the methodology of political inquiry*. Homewood, IL: Dorsey Press.
- Johnson, J. B., Reynolds, H. T., & Mycoff, J. D. (2007). *Political science research methods* (6th ed.). Washington, DC: CQ Press.
- Kinder, D. R., & Palfrey, T. R. (1993). *Experimental foundations of political science*. Ann Arbor: University of Michigan Press.
- King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Leege, D. C., & Francis, W. L. (1974). *Political research: Design, measurement, and analysis*. New York: Basic Books.
- Levy, P. S., & Lemeshow, S. (2003). *Sampling of populations: Methods and applications* (3rd ed.). New York: Wiley.
- Lohr, S. L. (1998). *Sampling: Design and analysis*. Pacific Grove, CA: Duxbury Press.
- Mack, N., Woodsong, C., MacQueen, K. M., Guest, G., & Namey, E. (2005). *Qualitative research methods: A data collector's field guide*. Research Triangle Park, NC: Family Health International.
- Manheim, J. B., Rich, R. C., Willnat, L., & Brians, C. L. (2007). *Empirical political analysis: Quantitative and qualitative research methods* (7th ed.). Essex, UK: Longman.
- Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

- McNabb, D. E. (2009). *Research methods for political science: Quantitative and qualitative methods* (2nd ed.). Armonk, NY: M. E. Sharpe.
- Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design and social measurement* (6th ed.). Thousand Oaks, CA: Sage.
- Ott, L. R., & Longnecker, M. T. (2001). *An introduction to statistical methods and data analysis*. Pacific Grove, CA: Brooks/Cole.
- Pollock, P., III. (2008). *The essentials of political analysis*. Washington, DC: CQ Press.
- Salmon, W. (1998). *Causality and explanation*. New York: Oxford University Press.
- Shiveley, W. P. (1990). *The craft of political research*. Englewood Cliffs, NJ: Prentice Hall.
- Silverman, D. (2001). *Interpreting qualitative data: Methods for analyzing talk, text, and interaction* (2nd ed.). Thousand Oaks, CA: Sage.
- Singleton, R. A., Jr., & Straits, B. C. (2004). *Approaches to social research*. New York: Oxford University Press.
- Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3rd ed.). Thousand Oaks, CA: Sage.

61

SURVEY RESEARCH

H. WHITT KILBURN

Grand Valley State University

Survey research is a major tool for bringing facts—data—to bear on political science theories (Brady, 2000). The way in which survey researchers do so, by collecting data from the few to generalize to the many, is once again undergoing a period of profound change. In the last significant period of change, survey research shifted from a reliance on face-to-face interviewing in respondent homes during the 1960s to the cheaper and faster world of telephone surveying in the 1970s and 1980s. Today, as the 21st century reaches its second decade, this transition toward a technology-mediated experience of the survey interview continues (Dillman, Smyth, & Christian, 2009). The revolution in digital communications technology has brought about even bigger changes, from the steady replacement of landlines with cellular phones to the expansion and habitual reliance of an ever-larger number of Americans on the Internet. And although survey researchers have dealt with public skepticism of polling and a refusal to participate before, today it is higher than ever. Nevertheless, survey research has always been an investigative tool shifting with the prevailing social trends (Tourangeau, 2004). As the study of survey research has become a scientific discipline of its own, survey research in political science is well prepared to meet these challenges and will adapt to do so.

This book, *21st Century Political Science: A Reference Handbook*, presents an appropriate opportunity

to take stock of the most important changes and sources of continuity in survey research. This chapter discusses four theses for the future. First, cellular telephone ownership has increased the number of households without a landline telephone, disrupting the traditional methodology for telephone surveys. Second, the web survey is already a major mode of survey research, and given the spread of broadband Internet access, its methodology will continue to develop. Third, these two changes have resulted in a blurring of what was once a fundamental distinction between surveys: the difference between probability and nonprobability sampling. And fourth, highlighting the place of survey research in a globalized world, cross-cultural (or cross-national) survey research will continue to open up new research opportunities. Prior to these ideas, however, to see where survey research heads into the future, this chapter must briefly review a few critical ideas about what a survey is, how survey questionnaires are written, and how survey sampling works.

Basic Ideas in Survey Research

Survey research in political science encompasses a great diversity of subject matter; the most well-known application is the mass opinion survey of an entire nation's voting-age population. Within the United States, the major political

science survey is the American National Election Studies (ANES); around the world, the World Values Survey is conducted in more than 40 countries. Across Europe, there are the long-standing Eurobarometer surveys and the more recent European Social Survey. These studies are only a few among the many in political science. (For others, see Chapter 94, “Voting Behavior,” in this handbook.)

Although these surveys all share a common concern for understanding the beliefs, attitudes, and values of mass democratic publics, a general definition of a scientific survey is surprisingly elusive, given the many ways in which survey research is conducted. At its core, survey research is the process of collecting data from a small part of a population to make general statements, or inferences, about characteristics of the larger population (de Leeuw, Hox, & Dillman, 2008a). These data are collected by having people answer questions to develop a set of systematic descriptions of the sample (Weisberg, 2005). The foundation of this process is built from writing a survey questionnaire and drawing a sample of individuals to interview.

Survey Questionnaire Design

In the early to mid-20th century, the work of writing survey questions and laying out a survey questionnaire resembled the art of a sculptor; aesthetic principles and skilled craft resulted in a work reflecting the vision of its creator and appreciated on her terms. Today, however, science has supplanted art; a robust literature testing practically all aspects of survey questionnaire design provides

strong guidance for researchers (Schaeffer & Presser, 2003). Although too voluminous to be summarized here, there are three essential aspects: (1) differences in question structure from open- to closed-ended, (2) how event recall and attitude questions are answered, and (3) consequences of the phrasing and order of survey questions.

Open and Closed Ended Questions

The most basic distinction within survey questionnaires is between open- and closed-ended questions. Open-ended questions are conversational, such as a question the ANES (2004) has asked about parties: “I’d now like to ask you what you think are the good and bad points about the two national parties. Is there anything in particular that you like about the Democratic Party?” If the respondent says, “Yes,” then the interviewer asks, “What is that?” The interviewer records everything the respondent says. Usually, survey researchers analyze open-ended responses by categorizing phrases, counting mentions of a theme such as *their handling of the economy*. Open-ended questions are very costly in survey time to administer and analyze by researchers; for these reasons, closed-ended questions are much more common in survey research. Closed-ended questions provide response options for a respondent to identify and select. There are many ways of structuring the response alternatives for closed-ended questions, but the most common is to ask a respondent to select an item from a rating scale.

Two different types of rating scales appear in Table 61.1. The upper half of the figure displays a bipolar response

Bipolar
<p>“We hear a lot of talk these days about liberals and conservatives. Here is a 7 point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale, or haven’t you thought much about this?”</p>
Branching
<p>“Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?” If the response is “Republican” or “Democrat”: “Would you call yourself a strong Republican [or Democrat] or a not very strong Republican [or Democrat]?” If the response is “Independent” or other: “Do you think of yourself as closer to the Republican Party or to the Democratic Party?”</p>

Table 61.1 Bipolar and Branching Scales

NOTE: The table displays two survey questions from the ANES (2004), in two common closed-ended formats, a bipolar and a branching scale. Both are used to create 7-point response scales for strength and direction of ideological beliefs and party identification.

scale for gauging a respondent's ideological identification. The scale ranges from two poles, *extremely liberal* to *extremely conservative*. In answering the question, the respondent is asked to select both a direction and strength of ideological identity. For the process to be as valid and reliable as possible, all points on the scale are labeled. Respondents who haven't thought much about it are excluded from the scale. In the lower half of Table 61.1, respondents are given a branching scale that unfolds in two steps. Respondents first select a direction of their party identification (Democrat, Republican, or Independent), then second, select a strength of identification, strong or not very strong, and if Independent, whether they lean toward either party. Although both scale types are acceptable, branching formats are preferred to bipolar response scales in which not all points on the bipolar scale are labeled (Krosnick & Berent, 1993). Telephone surveyors tend to rely on branching scales rather than verbal descriptions of bipolar response scales.

Answering Factual Recall and Attitude Questions

Closed-ended questions may be used to gauge respondent recall of objective factual information. For example, a researcher may ask, "How many days in the past week did you watch the national network news on TV?" The response alternatives range from "1 day" to "7 days." In answering such questions, respondents may simply attempt to recall the past week's schedule and estimate (or recall from working memory) the correct answer. (This description simplifies a complex process described in greater detail in Tourangeau, Rips, & Rasinski, 2000.) In political surveys, such questions are usually of less interest than those assessing a respondent's attitude or their positive or negative opinion about a political party, elected official, or policy issue. Although slightly different, the personal beliefs in Table 61.1 also contain an evaluative aspect and are similar to attitudes. For respondents, the process of answering attitude questions is fundamentally different and more complex.

For example, an attitude question commonly asked in U.S. surveys is presidential approval. In many contemporary surveys, respondents are asked, "Do you approve or disapprove of the way Barack Obama is handling his job as president?" When respondents answer such questions, they do not pull preexisting opinions out of memory. Instead, opinions are constructed on the fly. People draw on general values, predispositions, and fragments of prior beliefs in expressing an opinion through a process termed *belief sampling* (Tourangeau et al., 2000).

In the case of President Obama, respondents could have a wide range of positive and negative considerations brought to mind when asked. For example, a respondent may have recently seen a newscast describing bonuses paid to bailed-out bank executives on Wall

Street, and when the respondent is later asked about Obama, this consideration comes to mind, drawing him or her toward expressing disapproval of Obama's presidency. Generally, the balance of negative or positive considerations results in a similar direction of attitude expressed by the respondent. One implication of belief sampling is that researchers should not automatically include a "don't know" response alternative in survey questions based on their assumption that respondents do not have a preexisting opinion. Moreover, the importance of the concept of belief sampling is that it is useful for explaining context effects in surveys: how question wording and differences in question order can alter survey results. A more in-depth discussion of belief sampling and context effects is Schwarz, Knäuper, Oyserman, and Stich (2008), while Asher (2007) provides an accessible introduction with a clear application to the interpretation of poll results.

Survey Question Wording and Order

Clearly, question wording and order is important because of the influence of words on the sampling of considerations. Subtle changes in question wording can alter the considerations brought to mind. Asher (2007) discusses an example in which Americans were asked about their support for cuts in state services; when respondents were given the option of cutting what was termed *aid to the needy*, only 8% chose this option, but when the option was substituted with the term *public welfare*, 39% chose it. The term *welfare* apparently primed much more negative considerations. The most general advice about writing survey questions is to be aware of how such wording changes can affect results. Beyond that, questions should be worded in such a way that terms are defined as concretely as possible, using unbiased, simple language, while avoiding so-called double-barreled questions that refer to two subjects at once.

The order of survey questions can affect individual responses because particular questions prime certain considerations. For example, if one asks respondents if they would say that over the past year the nation's economy has gotten better, stayed about the same, or gotten worse, prior to a question on presidential approval, it would lead respondents to evaluate the president in light of economic considerations. Because of the potential for the order of questions to influence individual survey responses, wherever possible professional survey researchers randomize the order of questions.

At a practical level, when constructing a questionnaire researchers should have on hand a comprehensive reference. Authoritative, up-to-date textbooks are Fowler (2009) and Groves et al. (2004), and a slightly dated textbook is Weisberg, Krosnick, and Bowen (1996). A classic reference with practical advice is Dillman et al. (2009).

Random Sampling

Given a survey questionnaire and a population of individuals to study, the researcher must draw a sample from that population. Randomization is the cornerstone of probability sampling methods and is the professional standard for survey research. It is illustrated in its purest form in the simple random sample via the use of a sampling frame. A frame is a list of each and every individual within the population of interest. The sample will be drawn from the individuals on the frame; the frame should exactly mirror the population of interest, or else the sample will be subject to coverage error: the difference between the individuals appearing on the frame and in the population. In a simple random sample, all individuals in the population have a known and nonzero, equal chance of being selected. From the listed elements of the sampling frame, a random-number generator could be used to select the corresponding individuals listed sequentially on the frame.

Next, a researcher would administer a survey questionnaire. With answers to these questions recorded from each member of the sample, classical theories of statistical inference from any introductory statistics textbook could estimate population characteristics within a margin of error. With approximately 1,500 interviews, a researcher could estimate a characteristic (such as party identification) with a remarkable degree of precise accuracy, within roughly two percentage points of the true population value.

In sampling, however, there is often a disjuncture between the basic theory and feasible practice. Simple random samples are rarely ever applied, least of all for any survey of a geographically dispersed population, such as an entire region or nation. Consider two examples, the first from survey research involving face-to-face interviews of nationally representative American voters and the second from random digit dial (RDD) telephone surveying, the two polling methods most frequently used in large-scale research over the past century.

Face to Face Interviews and Area Probability Cluster Sampling

During the 1950s, the time characterized by Weisberg (2005) as the period of professionalization and expansion of survey research, a large proportion of American households still did not have a telephone. Major national surveys of the day such as the Gallup Poll and the ANES were conducted in person; interviewers traveled to the homes of survey respondents. Without an accurate sampling frame of American citizens of voting age (the ANES study population), a simple random sample was (and remains) impossible. Even if it were possible, meeting face-to-face with a simple random sample of Americans spread across the United States would be prohibitively expensive. So the method for drawing samples for such interviews has relied on an alternative that does not require a national sampling

frame and randomly distributed interviews: area probability cluster sampling (Weisberg et al., 1996).

Area probability cluster sampling for face-to-face surveys such as the ANES occurs in stages, resulting in a nationally representative sample of individuals interviewed in regional clusters. First, the United States would be divided into mutually exclusive primary sampling units (PSUs), such as metropolitan statistical areas (big cities) or sets of rural counties. Sampling is stratified by region, sampling units from within the north, south, east, and west of the country. Second, using maps of the areas within each PSU, multiple city blocks and similar rural areas are sampled as chunks. At the third stage, a sampling frame is constructed of all the housing units within these blocks. Then, fourth, individual housing units are sampled, and from within each one, an individual household member is selected for the interview. (A common method of selection is to interview the eligible person with the most recent birthday.) The result is a nationally representative sample; what may not be possible, however, is inference to each U.S. state, since interviews in small states may occur in only one city or county. There are many other aspects to this methodology; see Fowler (2009) for a further discussion.

This methodology has remained the platinum standard for achieving high-quality data and response rates and can be applied to any geographically defined area. In developing countries, it is the primary tool for national survey research. Nevertheless, face-to-face surveys are increasingly expensive to conduct. A current rule of thumb is that in the United States, the surveys cost approximately \$1,000 per interviewed respondent. Compared with approximately \$5 for 15 minutes of time with a respondent on an RDD poll, national face-to-face interview polls are cost prohibitive for nearly all purposes.

Telephone Surveys in the RDD Poll

Surveys conducted via the landline telephone became the norm in mass survey research with the advent of RDD methodology. Beginning in the 1970s, with sufficient coverage of landline telephones across the United States, the cheaper cost of contacting respondents over the phone meant that researchers could more quickly and efficiently complete their work. (For a discussion of these developments, see the 2007 special issue of *Public Opinion Quarterly* on telephone surveying.) Thus began the heyday of the RDD poll, which extended through the 1980s and into the 1990s.

To make telephone sampling cost-efficient, researchers had to determine how to isolate residential telephone numbers out of all the possible numbers, such as those no longer working or assigned to an address other than a residential household, which was the case for most telephone numbers. Randomly dialing any telephone number, similar to a simple random sample, would prove to be too

inefficient. The Mitofsky–Waksberg (MW) method was the first protocol to develop an efficient alternative, based on the logic of cluster sampling (Tourangeau, 2004). Residential telephone numbers have never been spread randomly throughout each combination of 7-digit telephone numbers; the MW methodology provided a way to reach more residential telephone numbers with less dialing.

The MW method is a two-stage sampling design, beginning with identification of sampling clusters and then continuing with sampling of individuals (Brick & Tucker, 2007). Consider a telephone number, given an area code (abbreviated AAA), a prefix (PPP), and the last four numbers (SSRR): AAA-PPP-SSRR. Because residential telephone numbers tend to be clustered together, in the WM method, numbers are selected in clusters. From an area code and prefix (AAA-PPP) bank, telephone numbers are sampled by randomly selecting an SS portion of the suffix. Then, in the second stage, two RR numbers are randomly chosen, and that telephone number is dialed. If it is not a residential number, that particular telephone bank is discarded, and the researcher moves on to the next cluster. If it is a residential number, then the researcher continues to dial additional numbers by randomly choosing RR numbers within the cluster, perhaps conducting as many as 10 interviews within the cluster.

The WM method is still used today (Brick & Tucker, 2007; Lepkowski et al., 2008). Of course, the method has evolved (Dillman et al., 2009), since over the past 20 years telephone service has expanded dramatically and residential lines are less densely assigned to particular banks of numbers, reducing the efficiency of RDD telephone surveys (Tucker & Lepkowski, 2008). As a result, today, researchers increasingly use RDD methods supplemented by purchased lists of working banks that function much like a sampling frame (Tourangeau, 2004).

Thus, RDD polls have become more expensive, particularly with the widespread use of caller identification and voicemail for screening calls from pollsters. Faced with increasingly difficult-to-reach respondents, some researchers may be tempted to take the cheaper, incorrect decision to replace these households in the sample with an additional, easier-to-reach household. Such a decision can cause problems when there is a significant difference between the two sets of respondents. For example, during the 2008 Democratic presidential primary in New Hampshire, survey researchers overestimated support for Senator Obama while underestimating it for Senator Clinton; the harder-to-reach households tended to support Clinton, skewing the results since they were excluded from samples (American Association for Public Opinion Research, 2009). Yet the bigger concern for the future of telephone surveys is the growth of cellular telephones and the potential for a coverage bias due to the increasing number of Americans who own cellular telephones but not landlines

and are thus excluded from (not covered in) the traditional telephone methodologies.

The Future of the RDD Poll and Cellular Phones

Standard estimates for the proportion of American households with only cellular telephone service (no landline phone) are based on the National Health Interview Survey, which began collecting data in 2003 on the implications of cellular telephone ownership for survey research. In the second half of 2008, about 20% of American households fit within this category. The proportion of cell-only households has been growing at a rate of approximately 2% to 3% every 6 months. Americans without landline telephone service are likely to be younger and poorer, and more likely to rent than to own homes or apartments. The largest wireless-only age-group is composed of persons from 25 to 29 years old, of whom approximately 40% live in wireless-only households, while among those aged 18 to 24, 33% do so (Blumberg & Luke, 2009).

The problem of undercoverage in traditional landline RDD polls is problematic to the extent that those persons systematically excluded from the surveys differ in relevant political characteristics from those who are not. Fortunately for political science, currently cell-only ownership appears to be largely unrelated to the political judgments asked in surveys. For instance, in the United States, among the youngest cohort of Americans, 18 to 25 years of age, although there appears to be little difference between cell-only young Americans and those with a landline on most measures of political beliefs and attitudes, those with only cell phones are significantly more likely to consume alcohol, among other behaviors that concern public health researchers (Blumberg, 2007; Keeter, Kennedy, Clark, Thompson, & Mokrzycki, 2007).

Yet despite the potential for a coverage bias in RDD polls affecting election surveys, for now, there appears to be little cause for immediate concern. Furthermore, any potential for coverage bias in RDD surveys is corrected, in the analysis of the data, by demographic weighting. Demographic sampling weights, calculated via a technique called *raking* or *sample balancing*, are typically required in the analysis of all contemporary RDD polls. Individual cases are weighted in an analysis to represent slightly less or greater than one observation to make the sample more closely match the known aggregate demographics of the population. For national surveys in the United States, surveys are usually matched to the U.S. Census Bureau's most recent Current Population Survey. For a technical discussion of weighting and its implications, see Biemer and Christ (2008) or Lepkowski et al. (2008).

It is unlikely that cellular telephones will be included routinely in mass surveys in the near future. Because cell

phone consumers in the United States must pay for incoming calls, the federal Telephone Consumer Protection Act prohibits anyone—survey researchers included—from using any autodialing technology to attempt to contact cell phone owners. (Often, telephone surveyors use automated technology to initiate the interviews, similar to that used by telemarketers.) Only under the circumstance in which the cell phone owner gave the researcher express prior consent to be contacted on the cell phone, or where a telephone interviewer manually dials the number, would such a contact be legal, thus increasing the cost of contacting cellular telephone owners. Generally, there currently exists no standard professional methodology for conducting cell phone surveys; the significantly higher costs in doing so mean that at least for the immediate future, in RDD polls, cellular telephones are more likely to be excluded from samples than to be included. Although cellular telephone surveys are out, web surveys are increasingly popular.

New Web Survey Methodology

Any casual web surfer must notice the ubiquity of surveys conducted over Internet web browsers, from the one- or two-question flash polls appearing on news websites to customer satisfaction surveys and e-mail solicitations for lengthy marketing studies. Web surveys are based on a wide range of survey sampling procedures; although news flash polls may not even be representative of visitors to a particular website, other web surveys are based on highly scientific, accurate probability samples. At its core, the web provides a survey mode: an interface for conducting surveys, independent of a particular sampling methodology. Although older survey methodologies, such as RDD polls, imply both the sampling method (RDD) and an interview mode (telephone), Web surveys imply nothing (Couper & Miller, 2008). The use of web surveys over a broad range of sampling purposes has spawned a rich array of developments in web survey technology, particularly software, the varieties of which are far too numerous and diffuse to effectively summarize here. (A comprehensive resource for technology and scholarship of web surveys is the website for Web Survey Methods: <http://www.websm.org>) Yet what is considered here are two concerns: One is a fundamental feature of web surveys and source of new research, and the second is an implication of an unfulfilled promise for web surveys to deliver cheap yet high-quality data.

Compared with survey methodologies for face-to-face or telephone interviews, web surveys are not interviewer assisted, since much survey research increasingly is through digital communications technology. But traditional paper-and-pencil surveys, too, are not interviewer assisted. And web surveys tend to be modeled after the paper-and-pencil surveys of generations ago; the layout of

the more popular online web survey companies offering free survey hosting are modeled after this layout. Some lament this state, calling for an investigation into the rich potential of web surveys to interact with the respondent through multimedia (Couper, 2007). Couper (2008) has written a comprehensive web survey design manual. Others find that the pencil-and-paper method has continued to be the most desirable for collecting reliable and valid data (Dillman et al., 2009).

As web survey methodologies developed, researchers hoped to uncover a new methodology that would lead to vast new methods for collecting new survey data. Yet survey research has not yet done so. At least for now, it does not appear that the administration of surveys via the web has delivered on hopes that this mode would recapture the response rates of yesteryear or of the telephone survey enterprise in its heyday. Response rates for web surveys appear to be no greater than those for other survey methods (Couper & Miller, 2008). Yet it does appear, however, that a web survey could complement other, more traditional survey modes as a way of reducing survey costs. Respondents can answer web surveys, while more expensive mail-back or face-to-face survey modes could be used for other respondents (Rookey, Hanway, & Dillman, 2008). The use of these mixed modes of survey research is likely to become a standard part of survey research projects (de Leeuw, Hox, & Dillman, 2008b).

One of the principal benefits of web surveys is the democratizing effect of Internet technology. Lower costs expand the reach of conducting surveys to more researchers, yet the familiar challenges from unequal Internet access still remain (Couper & Miller, 2008). Because of the continued digital divide (the systematic differences between those with home-based, consistent access to broadband Internet and those without), the challenge remains of respondent Internet access, representation, and generalizability from the sample to the population (Groves et al., 2004). Yet in response to the problem of coverage bias in web surveys, a number of innovations have emerged, blurring the distinction between probability and nonprobability sampling.

Blurring Probability and Nonprobability Sampling

One common type of web survey is the Internet panel, in which individuals volunteer to periodically complete web surveys produced by a researcher. These opt-in studies, which rely on volunteers recruited through web advertisements, ask for volunteers to join a firm's panel (a periodic survey) and periodically answer a few surveys. Although most popular with marketing and consumer research, some research firms conduct political polls in this form, such as YouGovPolimetrix. The survey methodology and its

application developed shortly after the turn of the 21st century, where in light of the decline of response rates for telephone surveys and in hopes of reaching younger respondents at relatively cheaper survey costs, web panels emerged (Dillman et al., 2009).

Yet panel surveys, while consisting of responses from an unrepresentative sample of Americans that complete the survey, are intended to facilitate generalizations to a more diverse population and sometimes even to the American public as a whole. When signing up to join a panel, respondents provide some basic demographic information about themselves, such as age, gender, race, formal educational attainment, and place of residence. Out of the people who volunteer and join the panel, a sample of these individuals is drawn to complete a survey, selected on the basis of these characteristics that should match those of the target population.

There are a number of potential sources of controversy with these surveys. From a more traditional, probability-sampling perspective, these surveys present a few potential difficulties: (a) coverage error—individuals without Internet access cannot participate and systematically differ from those who do; (b) self-selection bias—those panel members completing the survey are self-selected into it; (c) estimation of sampling error in the absence of randomization; and (d) panel-conditioning effects from repeatedly asking respondents to participate in surveys (Dillman et al., 2009). In accounting for these potential sources of error, Internet panel researchers have come up with a variety of technically sophisticated ways of analyzing the data to help the survey sample become more representative of the population of interest.

The potential for difference between the sample and population can be accounted for via sample weighting, in which the response of a particular individual is weighted to count for more or less than that one person's score, such that the aggregated individual responses are representative of a population. After the survey is administered, these weighting adjustments occur via a form of weighting called *propensity scoring*. Through propensity scoring, the demographic characteristics of a volunteer panel are weighted to more closely match the demographic characteristics of a similar, traditional probability-sample survey. Typically, researchers running an Internet panel will run a parallel, small-scale probability survey for the purpose of providing these data. For a discussion of propensity scoring and web surveys, see Lee (2006) and Lee and Valliant (2009).

Success of weighting to correct for bias depends on the similarity between the groups of persons included within the survey and those excluded. If the approximately 30% of Americans who do not have Internet access systematically differ from those who do, the weights may not successfully adjust for nonresponse. These studies depart from the traditional theory of probability-based sampling.

Yet the proponents of these surveys vigorously defend their methods as producing valid inferences.

Others, though, are not convinced and propose to view these studies as offering exploratory findings or as a vehicle for carrying out experimental tests (Dillman et al., 2009). The methodology of survey-based experimentation—designing an experimental manipulation within a survey—has developed strongly over the past decade. For an in-depth review of survey-based experimentation, see Lupia (2002), Sniderman and Grob (1996), and Gaines, Kuklinski, and Quirk (2007). Others see Internet panels as a return to the nonprobability polls (quota or convenience sampling) conducted prior to the institutionalization of probability sampling among professional pollsters (Bethlehem & Stoop, 2007). Despite criticism of the methodology and the declining response rates that now affect Internet panels as much as any other group, given the growth of Internet access and social networking websites (and web 2.0 technologies), panels will continue to grow. So although controversial, panels are likely to be increasingly popular among researchers.

There are probability-based alternatives for web surveys. Some panel providers have attempted to overcome the coverage problem of the Internet by recruiting participants through a probability-based method. For example, in the Netherlands, CentERdata recruited participants for their 2007 longitudinal Internet studies for the social sciences (LISS) panel through a probability sample of households drawn from national census records. Similarly, the most recent ANES 2008 presidential election survey included a monthly web survey in which participants were recruited to participate through an RDD sample. In both examples, if respondents did not already have a personal computer and Internet access, they were given one to use during the duration of the study. Of course, questions remain about the conditioning effect of computer access and web surveys; relying on RDD sampling with the rise of cellular telephones brings about its own challenges. There remains much to be learned about the subject.

Cross-National Surveying

An area of major concern in survey research is the development of survey research methods conducted across linguistic, national, or even continental boundaries. The development of valid, reliable survey research methodologies (called *cross cultural* methods) is a major theme for the 21st century (de Leeuw et al., 2008b). Cross-national survey research has opened up many new research frontiers in the comparative study of public opinion; methodologically, survey researchers continue to investigate methods of standardizing survey administration to facilitate cross-national comparisons. Authoritative

texts on the subject are Harkness, Van de Vijver, and Mohler (2003) and de Leeuw et al. (2008b).

Conclusion

The challenges facing survey research for the 21st century are great, but it would be premature to begin penning the method's obituary. Survey research has faced similar challenges in the past and been the subject of criticism that the industry would decline. When survey researchers faced severely declining response rates for face-to-face surveys in the 1960s, some questioned whether surveys would survive. Instead of declining, survey research thrived, leading to the widespread scientific study of survey questionnaires and sampling methodologies. So today, survey researchers will meet the challenges of the current era. No other research tool facilitates the study of population characteristics on the basis of a relatively small sample as well as survey research. Even challenges such as the development of cellular telephone technologies will likely prove to be surmountable. The spread of Internet access, further development of social networking technology, and the continued growth of exclusive cell phone ownership will likely be future research subjects, ensuring the place of survey research in political science.

References and Further Readings

- American Association for Public Opinion Research, Ad Hoc Committee on the 2008 Presidential Primary Polling. (2009). *An evaluation of the methodology of the 2008 pre election primary polls*. Retrieved May 1, 2009, from <http://www.aapor.org>
- American National Election Studies. (2004). *Codebook variable documentation file*. Retrieved April 9, 2010, from <http://www.electionstudies.org/study/pages/2004prepost/nes04var.txt>
- Asher, H. B. (2007). *Polling and the public: What every citizen should know*. Washington, DC: CQ Press.
- Bethlehem, J., & Stoop, I. (2007). *Online panels A paradigm theft?* Paper presented at the meeting of the Association for Survey Computing, University of Southampton, UK.
- Biemer, P. P., & Christ, S. (2008). Weighting survey data. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 317-341). New York: Lawrence Erlbaum.
- Blumberg, S. J. (2007). Coverage bias in traditional telephone surveys of low income and young adults. *Public Opinion Quarterly*, 71, 734-749.
- Blumberg, S. J., & Luke, J. V. (2009). *Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2008*. National Center for Health Statistics. Retrieved May 1, 2009, from <http://www.cdc.gov/nchs/nhis.htm>
- Brady, H. E. (2000). Contributions of survey research to political science. *PS: Political Science and Politics*, 33, 47-57.
- Brick, J. M., & Tucker, C. (2007). Mitofsky Waksberg: Learning from the past. *Public Opinion Quarterly*, 71, 703-716.
- Couper, M. P. (2007). *Whither the web: Web 2.0 and the changing world of web surveys*. Paper presented at the meeting of the Association for Survey Computing, University of Southampton, UK.
- Couper, M. P. (2008). *Designing effective web surveys*. New York: Cambridge University Press.
- Couper, M. P., & Miller, P. V. (2008). Web survey methods: Introduction. *Public Opinion Quarterly*, 72, 831-835.
- de Leeuw, E. D., Hox, J. J., & Dillman, D. A. (2008a). The cornerstones of survey research. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 1-17). New York: Lawrence Erlbaum.
- de Leeuw, E. D., Hox, J. J., & Dillman, D. A. (2008b). *International handbook of survey methodology*. New York: Lawrence Erlbaum.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: Wiley.
- Fowler, F. J., Jr. (2009). *Survey research methods*. Thousand Oaks, CA: Sage.
- Gaines, B. J., Kuklinski, J. H., & Quirk, P. J. (2007). The logic of the survey experiment reexamined. *Political Analysis*, 15, 1-20.
- Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: Wiley.
- Harkness, J. A., Van de Vijver, F. J. R., & Mohler, P. P. (2003). *Cross cultural survey methods*. Hoboken, NJ: Wiley.
- Keeter, S., Kennedy, C., Clark, A., Tompson, T., & Mokrzycki, M. (2007). What's missing from national landline RDD surveys? The impact of the growing cell only population. *Public Opinion Quarterly*, 71, 772-792.
- Kempf, A. M., & Remington, P. L. (2007). New challenges for telephone survey research in the twenty first century. *Annual Review of Public Health*, 28, 113-126.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50, 537-567.
- Krosnick, J. A., & Berent, M. K. (1993). Comparisons of party identification and policy preferences: The impact of survey question format. *American Journal of Political Science*, 37, 941-964.
- Lavrakas, P. J. (1993). *Telephone survey methods: Sampling, selection, and supervision*. Newbury Park, CA: Sage.
- Lee, S. (2006). Propensity score adjustment as a weighting scheme for volunteer panel web surveys. *Journal of Official Statistics*, 22, 329-349.
- Lee, S., & Valliant, R. (2009). Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods and Research*, 37, 319-343.
- Lepkowski, J. M., Tucker, C., Brick, J. M., de Leeuw, E. D., Japac, L., Lavrakas, P. J., et al. (Eds.). (2008). *Advances in telephone survey methodology*. Hoboken, NJ: Wiley.
- Lupia, A. (2002). New ideas in experimental political science. *Political Analysis*, 10, 319-324.

- Rookey, B. D., Hanway, S., & Dillman, D. A. (2008). Does a probability based household panel benefit from assignment to postal response as an alternative to Internet only? *Public Opinion Quarterly*, 72, 962-984.
- Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, 29, 65-88.
- Schwarz, N., Knäuper, B., Oyserman, D., & Stich, C. (2008). The psychology of asking questions. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 18-34). New York: Lawrence Erlbaum.
- Sniderman, P. M., & Grob, D. B. (1996). Innovations in experimental design in attitude surveys. *Annual Review of Sociology*, 22, 377-399.
- Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, 55, 775-801.
- Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The psychology of the survey response*. New York: Cambridge University Press.
- Tucker, C., & Lepkowski, J. M. (2008). Telephone survey methods: Adapting to change. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, et al. (Eds.), *Advances in telephone survey methodology* (pp. 3-28). Hoboken, NJ: Wiley.
- Web survey methods: <http://www.websm.org>
- Weisberg, H. F. (2005). *The total survey error approach: A guide to the new science of survey research*. Chicago: University of Chicago Press.
- Weisberg, H., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3rd ed.). Thousand Oaks, CA: Sage.

62

EXPERIMENTS

ELIZABETH A. BENNION

Indiana University South Bend

Experimental research experienced a resurgence in the 21st century. This resurgence was led by a group of scholars at Yale University who persuasively argued that randomized intervention into real-world settings should “occupy a central place in political science” (Green & Gerber, 2002, p. 808). Committed to the belief that the value of survey research had been overstated and the value of field experiments was underappreciated, they set out to explore and promote the “untapped potential of field experiments” (p. 808). Working through Yale’s Institution for Social and Policy Studies, Green and Gerber set up a summer workshop on field experiments, inviting social scientists across the nation (and world) to join them in this shared endeavor. Meanwhile, they trained their graduate students to conduct field experiments, inspiring a series of doctoral dissertations and academic articles using field experimentation. This chapter discusses the experimental method, compares the experimental method to survey-based research, and stresses the importance of random assignment of experimental treatments. The chapter also explains the difference between laboratory experiments and field experiments, highlights the wide range of applications for experimental studies, and briefly discusses the policy implications and future directions of experimental research in political science.

Theory

Most chapters in this handbook discuss substantive, topic-based areas of the discipline. These subfields are driven by assumptions, or theories, about the way the political world works. In contrast, this chapter focuses on a specific method for studying political phenomena: the experimental method. This method is designed to test substantive theories about the empirical world. Experiments are based on the assumption that political scientists can investigate the political world by designing specific interventions that change political behavior or policy outcomes in measurable ways.

The Experimental Method

An experiment is a method used to study cause and effect. The point is to examine the relationship between two or more variables. A variable refers to a measurable attribute (e.g., age, sex, educational attainment, or partisanship) that varies over time or among individuals. Experiments involve the deliberate manipulation of one variable, while trying to hold all other variables constant. By changing one variable while measuring another, the experimental method allows researchers to draw conclusions about cause and effect with far more certainty than any nonexperimental method. The variable manipulated by the researcher is

called the *independent variable* while the *dependent variable* is the change in behavior measured by the researcher. The logic is clear: If the independent variable is the only thing that is changed, then the independent variable is responsible for any change in the dependent variable. All other variables that might affect the results are called *confounding variables*. By carefully assigning subjects to treatment and control groups, researchers can ensure that confounding variables are evenly distributed among participants in both groups so that the effect of the experimental treatment itself can be isolated and measured.

Treatment and Control Groups

To conduct an experiment, the researcher divides research subjects (sometimes called participants) into a control group and a treatment group. The control group receives no treatment, while the treatment group receives a specific intervention. Suppose a political scientist wants to investigate whether calling people and reminding them to vote will actually increase the likelihood that they will cast a vote on election day. The phone call is the independent variable. The researcher wishes to determine whether phone calls can increase voter turnout. The dependent variable is voter turnout. Voter turnout records can be obtained from the county clerk or secretary of state. Turnout is the dependent variable because the researcher's hypothesis is that performance on this variable (level of turnout) depends on the independent variable (whether the person received a reminder phone call). To test the effectiveness of the phone calls, members of the control group do not receive a reminder call before election day, while members of the treatment group receive the reminder call. The researcher expects that people in the treatment group will, on average, be more likely to vote than people in the control group. The experiment allows this hypothesis to be tested empirically.

Confounding variables in the example given might include age, sex, partisanship, educational attainment, political interest, and past voter history. Each of these factors is correlated with voter turnout. Older people, women, strong partisans, and educated citizens with an interest in politics are more likely to vote than their counterparts. The best predictor of whether people will vote is their past behavior. People who have voted in the past are most likely to vote in the future. These factors may be more important than the reminder call in determining whether an individual will vote on election day. In a laboratory setting, researchers often match the treatment and control groups according to relevant characteristics to reduce the effects of confounding variables. An even better alternative is to randomly select participants into the treatment or control group.

Random Assignment

A random selection process ensures that every subject has an equal chance of being selected into the treatment group. As a check, the researcher can compare the characteristics of the treatment and control groups to assure readers that the groups really are similar along all relevant dimensions. Rather than drawing names out of a hat or flipping a coin, political scientists in the 21st century use computers to assign subjects to a treatment or control group. Using a random number generator to assign half of the participants to a treatment group and the other half to a control group ensures that the treatment group and the control group do not differ in terms of their politically relevant characteristics. This technique is particularly useful when the number of participants is large. The larger the number of participants, the less likely it is that members of a treatment group share some unidentified behavior-changing characteristic that could affect their performance on the dependent variable.

There is an added benefit of random selection for experiments conducted outside the laboratory. The random selection assures that other stimuli that might affect participant behavior (e.g., candidate mailings, campaign commercials, and television stories) will reach both the treatment and control groups. This means any effects of such outside stimuli will shape behavior of both the treatment and control group in roughly equal (though immeasurable) ways. In contrast, the random assignment of the researcher's intervention assures that only a representative sample of participants will receive the treatment of interest. In this way, researchers can isolate the specific influence of the intervention they are studying.

Comparing Different Interventions

There are many variations of this basic experimental method. The most common are comparisons of different treatments and the use of a placebo group. To test multiple treatments, a researcher simply creates additional treatment groups. The researcher in the earlier example may want to know whether phone calls or door hangers are more effective in getting people to vote on election day. The researcher might assign one third of the registered voters in a precinct to the control group, another third to Treatment Group A, and the final third to Treatment Group B. The control group would not receive any reminders. Treatment Group A would receive a reminder phone call encouraging them to vote. Treatment Group B would receive a door hanger reminding them to vote on election day. As long as the researcher uses random selection to assign subjects to one of the three groups, this study will effectively compare the relative impact of making a phone call versus leaving a written message on a prospective voter's door. By comparing the turnout rates of subjects in each of the three groups, the researcher can determine which approach is

most effective at getting people to the polls. Similarly, an experiment might compare the relative effectiveness of two different phone call scripts or two different door hangers to see which message is most effective in getting people to participate in the electoral process.

Understanding the Placebo Effect

Sometimes political scientists are concerned that they are measuring the effect of intervening in people's lives, rather than the effect of a specific treatment. To address this concern, one could use a placebo group. To understand this approach, it is helpful to think about the field of medicine.

Medical studies frequently employ the use of placebo groups to disentangle the psychological effects of receiving treatment from the actual physiological effects of the treatment itself. A placebo is a phony medical intervention that leads the recipient to believe that his or her medical condition may be improved. One common placebo treatment is an inert sugar pill. Subjects in a clinical trial may be divided into three groups: a control group, a treatment group, and a placebo group. The control group receives no medicine. The treatment group receives the medicine being tested. The placebo group receives the (medically ineffective) sugar pill. Subjects do not know whether they have received the new wonder drug or the inert sugar pill. The placebo effect is well documented. People frequently report feeling better after treatment, even if they receive the placebo.

Although political scientists are less likely to use placebo groups than their colleagues in the field of medicine, some recent studies demonstrate the effective use of placebo groups in political research. One recent study used a placebo-controlled experiment targeting households with two registered voters (Nickerson, 2008). Residents who answered the door were exposed to either encouragement to get out and vote (treatment) or a recycling message (placebo). The placebo treatment was used to address the possibility that people who answer the door and talk to strangers also may have a higher propensity to vote. This could be true because they are more civic-minded or simply because they are alive, mobile, and still living in the voting precinct. The fact that the treatment group voted at higher rates than the control and placebo groups increases confidence that the get-out-the-vote treatment was effective.

Randomized Experiments Versus Survey Data

The vast majority of work in political science relies on nonexperimental data. Since the early 1950s, surveys have been the mainstay of political behavior research. Earlier political scientists conducted some controlled experiments¹ (e.g., Gosnell, 1927; Hartman, 1936–1937), but their work was seldom replicated. In the 1950s, as the principles of

probability sampling and survey research became better known, political scientists sought to offer complete explanations for political phenomena. Surveys seemed ideally suited to this task, allowing researchers to take into account a wide range of demographic, economic, and social-psychological characteristics that shape political attitudes and behavior. In addition, surveys seemed better able to address big-picture questions of interest to political scientists, including topics like political culture, party identification, and support for the political system (Green & Gerber, 2002).

Surveys offer a relatively inexpensive way to study political attitudes and behavior from a nationally representative sample. A sample of only 1,000 Americans can provide a snapshot of public opinion that is highly accurate. With only 1,000 respondents, researchers can be 95% sure that they have captured public opinion with a mere $\pm 3\%$ margin of error. For example, if 57% of all survey respondents say that they approve of the job the president is doing in office, one can be 95% sure that the president's true level of support among citizens is between 54% and 60%.

Although survey data provide social scientists with valuable research opportunities, this approach has several inherent drawbacks. Consider the example of voter mobilization. A survey conducted using a randomly selected sample can provide a very good estimate of the percentage of people who voted, or at least the number of people nationwide who are likely to report voting, on election day. The problem is that this behavior is self-reported. By comparing survey data on voter turnout to official election records, political scientists have discovered that people overreport voting. People tell pollsters that they voted when they in fact did not. This is also true of other socially desirable behaviors. When people know that they should do something, they are more likely to report that they did it whether or not this report is accurate.

Voter turnout rates can be obtained through official election records. Political scientists wish to know not only how many people went to the polls on election day, but also why they voted while others failed to do so. To determine the most effective way to mobilize people to vote, survey researchers ask citizens to recall having been contacted by political campaigns. Unfortunately, the survey researcher must rely on respondents' self-reports that the contacts actually occurred. If voters are more likely than nonvoters to report campaign contact when none occurred, the analysis might overestimate the effect of contact on voter turnout. On the other hand, if nonvoters are more likely than voters to incorrectly report campaign contact, the analysis will underestimate the effectiveness of campaign contact. In addition to the problem of potential reporting bias, the nature of the contact between the campaign and the voter is usually unclear. It is difficult to distinguish between face-to-face contacts and phone contacts or between single contacts and multiple contacts. It is also difficult to determine what type of group contacted the

voter or what message was used. Survey questions are often not specific enough to detect these differences, and voter recollection is limited.

Even if these campaign contacts were accurately recorded, one could not be sure the contacts really increased turnout. It is possible that political campaigns targeted likely voters. Campaigns use state voter files, including voter history, to select voters to target. The correlation between contacts and turnout may simply reflect the fact that campaigns targeted likely voters. Experiments provide a more precise tool to isolate cause and effect.

Experiments isolate cause and effect by determining how a change in one variable causes change in another variable. Unlike survey researchers, experimenters know precisely what treatments each subject received. Most often, experimental research studies also allow the researcher to observe the actual outcome of the treatment. Neither the treatment nor the outcome is self-reported. In the case of a voter mobilization field experiment, for example, the researcher randomly assigns subjects to the treatment or control group and then delivers specific treatments (phone calls, face-to-face visits, etc.) to each subject, keeping careful records of who received the treatment. After election day, official voter records are examined to compare the voter turnout of the treatment group to that of the control group. Because of random assignment, the researcher knows that mobilization messengers did not target high-propensity or low-propensity voters. Because the study relies on actual records, rather than self-reports, the researcher need not worry about reporting bias. By comparing the turnout of the treatment and control groups, researchers can determine the precise effect of specific mobilization tactics.

Laboratory Experiments Versus Field Experiments

As political scientists recognized the limits of survey-based research, the late 20th century ushered in a renewed interest in experiments. Researchers began to modify surveys to include embedded experiments. Survey-based experiments were conducted by randomly assigning respondents to receive different versions of the same question to study how question content and wording shape people's answers to questions on politically sensitive topics such as racial attitudes (Hurwitz & Peffley, 1998; Sniderman & Grob, 1996). Using split samples is one way to avoid, or measure, the reporting bias that can undermine survey-based research. Researchers also designed laboratory experiments to study such topics as media exposure (Ansolabehere & Iyengar, 1995; Iyengar & Kinder, 1987; Iyengar, Peters, & Kinder, 1982), collective action (Dawes, Orbell, Simmons, & van de Kragt, 1986), and legislative bargaining (McKelvey & Ordeshook, 1990). Using human behavior laboratories, social scientists can set up lab experiments that are similar to those conducted by their colleagues in the physical sciences. The researcher creates

equivalent groups through matching or randomization and then follows one of three basic protocols: (1) administering a treatment to one group but not to the other, (2) administering the treatment to one group and a placebo treatment to the other group, or (3) administering different treatments to different groups.

Studying Challenging Topics Using Laboratory Experiments

Laboratory experiments can be useful in detecting prejudice, stereotyping, and other forms of bias that people might not readily admit to a person administering a survey. Such experiments are also useful at isolating specific variables that would be difficult, if not impossible, to isolate in the real world. For example, Sapiro (1991–1992) conducted an experiment on gender stereotypes. Interested in difficult-to-detect, perhaps nonconscious, forms of sexism, Sapiro showed students campaign speeches for hypothetical candidates for the U.S. House of Representatives. The text of the speech was taken from a real speech by a U.S. senator, selected because it provided little information about policy proposals, partisanship, or political ideology. The text for each candidate was identical, except that one was labeled “Speech by John Leeds” while the other version was labeled “Speech by Joan Leeds.” The students used gender stereotypes to determine how competent the candidate would be in solving specific kinds of problems. Students gave Joan higher competence ratings than John in specific policy areas, including improving our educational system, maintaining honesty and integrity in government, and dealing with health problems. None of these issues were directly mentioned in the speech. In contrast, students gave higher competence scores to John versus Joan when asked to rate the candidate's competence in the stereotypically male domains of dealing with military issues and making decisions on farm issues. Students were also less likely to think that female candidate Joan Leeds would win the election. Because the researcher so carefully controlled the experiment, offering students identical information except candidate name, the importance of candidate gender in shaping voters' perceptions is clearly demonstrated. Simply asking voters to answer questions about real-life male and female candidates would not prove that gender stereotypes are driving people's perceptions or responses. Real-life candidates have different personalities, ideologies, writing styles, speech patterns, and campaign strategies, all of which shape voter perceptions of the candidates.

The Limitations of Lab Based Experiments

The major question about lab-based experiments is whether they provide findings that will apply to the real world outside the laboratory. The results of laboratory experiments may not always be generalizable outside the lab. One potential problem is the requirement that all subjects

participating in a lab-based experiment must provide informed consent. Informed consent means that subjects are aware that they are being studied. There is a real concern that subjects may act differently because they know they are being watched. Researchers try to limit the effects of this potential problem by obscuring the specific research question and variable of interest, while accurately highlighting potential benefits and risks of participation. For example, researchers studying the effects of negative campaign ads might embed these ads within a newscast, telling subjects only that they are looking at selective perceptions of news programs (Ansolabehere & Iyengar, 1995). Even if the exact nature of the experiment is successfully obscured, the fact that so many political scientists use undergraduate students as research subjects reduces the likelihood that the findings are applicable to a full range of people outside the lab. It is unlikely that college students are representative of the population as a whole. Recognizing this limitation, some scholars have begun taking their experimental labs on the road. For example, Iyengar selected shopping malls as a laboratory for a series of experiments designed to test the effects of negative advertisements on respondents' knowledge, attitudes, and likelihood of voting. In this way, Iyengar was able to capitalize on the benefits of a controlled experiment while broadening his pool of subjects beyond the college campus (Ansolabehere & Iyengar).

Even if a more representative pool of participants is identified, there is still reason to doubt the generalizability of lab-based experiments. People make decisions based on a range of factors, including self-interest, rationality, and political ideology. However, human behavior is also shaped by the degree to which people believe their decisions will be scrutinized by others, the particular context in which a decision is made, and the manner in which participants are selected. If these circumstances do not reflect the real-world environment in which political decisions are made, the results may not be generalizable outside the lab. Despite these shortcomings, laboratory experiments can provide an important, even critical, first step to understanding people's political decision-making processes. They can also produce important findings that can later be tested, when possible, outside the laboratory.

Field Experiments: The Real World as Political Laboratory

Social scientists who wish to test hypotheses in the real world often turn to field experiments. Unlike laboratory experiments, field experiments examine an intervention or treatment in the real world, in naturally occurring environments. To maximize the realistic nature of the experiment, social scientists often use subjects who are unaware that they are participating in an experiment. Government requirements regarding the protection of human subjects require that subjects sign a consent form (or receive a study information sheet) unless the research design relies

solely on interactions that might take place anyway in the absence of the study and on public data available without the subjects' consent. As with laboratory experiments, researchers must also convince an institutional review board that the experiment will not in any way harm subjects and that the identity and confidentiality of all participants will be protected. Voter mobilization experiments meet all of the conditions for waiving informed consent. First, researchers are not doing anything other campaigns and political groups do not do in an election season. Second, lists of registered voters are collected from public voter files, and voting behavior is determined using these same public files. Finally, review boards are unlikely to argue that asking somebody to vote is likely to cause harm.

Natural Experiments: Exploiting As-If-Random Conditions

Sometimes political scientists conduct experiments without using random assignment. These studies, called *quasi experiments* or *natural experiments*, are conducted when real-life circumstances approximate the conditions of a randomized experiment. With quasi experiments, researchers observe differences between groups without assigning subjects to treatment and control groups or manipulating the treatment variable. Instead, researchers take advantage of a predetermined change, such as a new law or policy, designed to alter public behavior. For example, researchers might study the effects of a new gun control law by comparing homicide rates before and after implementation (Bogus, 1992). Studies like these that measure changes in the entire population reduce the problem of an unrepresentative treatment group by eliminating the possibility that people self-selected the treatment. Another approach is to select two different cities with comparable population sizes, education levels, racial and ethnic diversity, and pre-gun-ban crime rates and compare the homicide rates and gun-related crime rates after a ban was enacted in one city but not the other (Bogus, 1992). The key to making a convincing case would be to demonstrate that the two cities are, in fact, similar with regard to all characteristics that might affect the crime and homicide rates. Ideally, they would also have identical crime and homicide rates before the ban was put into place. The goal with natural experiments is to establish that the treatment and control groups will perform as if they were randomly selected.

Unfortunately, many social and policy changes do not meet this as-if-random requirement. For example, comparing the performance of students at a new magnet school to the performance of other public school students would not provide a good measure of the success of the new school in promoting student achievement. The fact that students and parents self-select into the magnet school may lead to higher performance among charter school children. Any differences in performance between children at the regular public school and those at the new magnet school may be due to selection bias. The magnet school may have attracted

high-achieving students with involved parents. Random assignment can solve this problem of selection bias. A lottery system would be desirable from a research point of view but is not always practical or desirable.

Quasi experiments are particularly useful when it is impossible for political scientists to control the variables of interest. Although political boundaries provide a popular basis for natural experiments, many other socially occurring phenomena may present the possibility for this kind of research design. As a first step in a government-funded series of experiments, Brady and McNulty (2004) used a natural experiment to explore how the costs of voting affect turnout. They studied California's special gubernatorial recall election of 2003, in which Arnold Schwarzenegger became governor. The elections supervisor in Los Angeles County consolidated the number of district voting precincts from 5,231 to 1,885. For some voters, the distance to their polling place was increased, while for others, it remained the same. The group that had to drive farther to get to the polls became the treatment group for a natural experiment studying how the costs of voting affect turnout. The key question is whether assignment of voters to polling places in the 2003 election was as-if-random with respect to characteristics that affect people's likelihood of voting. Did the county elections supervisor close some polling places and not others in ways that were correlated with this predisposition to vote? Brady and McNulty find some evidence for a lack of pretreatment equivalence between groups of voters who had their polling place changed (i.e., the treatment group) and those who did not. This threatens the validity of their findings. Fortunately, in this case, the pretreatment differences between the groups are small, relative to the reduction in turnout associated with increased voting costs. This indicates a strong likelihood that forcing people to drive farther to get to a polling place reduces turnout on election day. Scholars have begun to evaluate the plausibility of various natural experiments in the social sciences based on the degree to which they meet the as-if-random requirement (Dunning, 2008).

Applications

Students of politics are filled with questions about why and how politics works. Experiments conducted in the lab, in the field, or embedded within surveys can further our understanding of the political world. Applications are numerous, but three areas of investigation have been particularly likely to generate experimental research: negative advertising, voter mobilization, and racial attitudes.

Lab Experiments on Negative Advertising

Scholars have long debated the effects of negative political advertising. Conventional wisdom holds that people dislike the ads but that they work. Most of this scholarship

went unnoticed by the media and political consultants until Ansolabehere and Iyengar (1995) published *Going Negative: How Attack Ads Shrink and Polarize the Electorate*. Based on lab experiments and observations of U.S. political campaigns, the authors argue that negative advertising depresses voter turnout and that political consultants intentionally use ads for this purpose. The authors suggest that negative ads work better for Republicans than for Democrats and better for men than for women and also that negative ads work better than positive ones. They caution that as independent voters are driven away by negativity, the voting public is reduced to its partisan extremes. A 1996 study challenged these conclusions, finding that negative ads can promote political participation, especially among uninformed voters (Wattenberg & Brians, 1996). Using survey data, the authors found that citizens who report being exposed to negative ads are more likely to vote than those who do not comment on such ads. They argue that Ansolabehere and Iyengar's findings must not apply outside the lab. Given this high-profile dispute, several political scientists conducted a review of the literature on the topic and ultimately concluded that there is little evidence that negative advertisements are especially disliked, more effective than positive ads, or detrimental to participation in the electoral process (Lau, Sigelman, Heldman, & Babbitt, 1999). Political consultants remained convinced that negative advertising works, leading to a flurry of 21st-century experiments testing the effects of negative ads in a variety of forms and contexts.

Field Experiments on Voter Mobilization

Voter mobilization studies have been the subject of a variety of natural and randomized field experiments. The ability to test specific mobilization techniques, to accurately record the treatments received, and to bypass the problem of self-reporting using official voting records makes the experimental method ideal for this line of research. Scholars and practitioners of the art of campaigning have devoted significant attention to randomized field experimentation since Gerber and Green's (2000) article reporting on the effectiveness of different voter mobilization techniques. The success of this enterprise is documented in the 2004 release (and 2008 second edition) of Green and Gerber's (2008) instant hit, *Get Out the Vote! How to Increase Voter Turnout*. Targeting academic researchers and political practitioners, the book summarizes the results of dozens of voter mobilization field experiments conducted and published since the turn of the century. Scholars working in this area were also invited to publish their completed studies in a special edition of the *Annals of the American Academy of Political and Social Science*, titled "The Science of Voter Mobilization" (Green & Gerber, 2005). The journal featured a collection of articles by political scientists using randomized field experiments to test the effectiveness of different voter

mobilization methods, messages, and canvassers in a wide range of contexts.

Survey-Based Experiments on Racial Attitudes

Racial attitudes are difficult to study because of most people's reluctance to admit prejudice. Survey-based experiments have proven an unobtrusive way to measure racial attitudes and the effects of these attitudes on popular support for various government policies. Survey-based experiments challenge previous survey data that suggested whites in the South resemble the rest of the country in their racial attitudes (Kuklinski, Cobb, & Gilens, 1997). The authors of these experiments argue that previous survey results were contaminated by social desirability. By randomly assigning respondents to different forms of the question, cuing or not cuing people to think about race, political scientists can get around this problem. For example, a study of the effects of racial attitudes might ask people's impressions of a welfare recipient described as either a white or a black woman in her early 30s with a 10-year-old child who has been on welfare for the past year. How likely is it that she will have more children to get a bigger welfare check? How likely is it that she will look for a job? Because subjects were randomly assigned to receive either the black or the white version of the (otherwise identical) question, researchers can measure the effect of racial attitudes without directly asking the respondent to compare whites with blacks (Gilens, 1999). Based on these experiments, political scientists have argued that racial attitudes dominate public perceptions of welfare, with black stereotypes predicting much of the opposition to welfare programs (Gilens, 1999). In contrast, it is clear that public opposition to affirmative action is driven less by racial prejudice than commonly believed (Kuklinski et al., 1997). Political scientists continue to use experiments to investigate the effects of racial attitudes on political identity, attitudes, and behavior.

Other Uses

Experimental methodology has broad application to questions about the effectiveness of a wide range of social interventions. Although political scientists initially viewed the random assignment of social interventions in real-world settings (outside medicine) as impractical, the use of field experiments is gaining popularity and encouraging collaborations among scholars from many disciplines interested in political questions. Sage Publications published a special issue of the *American Behavioral Scientist* titled "Field Experiments in the Political Sciences" (Green & Gerber, 2004). This work crosses the disciplines of political science, social psychology, social work, criminology, and public policy. Topics include the relationship between campaign spending and electoral victory, how to frame messages to get patients to seek preventative care,

how to evaluate the effectiveness of social welfare program reforms, the difficulty of evaluating crime prevention programs, and the effectiveness of school voucher programs on academic performance.

Limitations and the Need for Replication

Although experiments excel at testing causal relationships, they are not without limitations. There are several common criticisms that have limited the use of randomized experiments in the discipline. Each criticism reflects valid concerns, but critics often overstate the extent of these limitations. Large experiments can require a great deal of time and money, but other forms of research are also time-, labor-, and (sometimes) capital-intensive. Moreover, many of the costs of field experiments can be covered by political organizations, agencies, and foundations that hire academics to evaluate their efforts.

Some critics correctly argue that experiments may produce contradictory findings. Whether because of sampling error or differences in experimental design, experiments may produce incompatible findings. However, this is true of studies based on a wide range of data collection methods. Clear, detailed descriptions of the experimental protocol and further experimentation and replication can help resolve these issues.

Other critics argue that experimental research frequently fails to offer a clear explanation for why a specific intervention produced a specific effect. This shortcoming can also be solved through additional experimentation. Researchers can vary the stimulus to determine which aspect of the treatment is producing the demonstrated effect. Researchers can also measure variables that are thought to affect the relationship between the intervention and the dependent variable.

In addition, critics argue that experimental results might not be applicable to the real world of politics. Replication in different contexts, including field experiments outside the lab, can boost confidence in the external validity of experimental findings. Each of these criticisms, while valid, points to the need for additional experimentation.

Finally, randomized experiments are sometimes dismissed as impractical, either because the subjects of investigation are too broad and complex or because it is thought that key political actors cannot (or will not) randomly assign groups to different types of interventions. It is true that political scientists cannot randomly assign states or countries to different forms of government, legal systems, or public policies. Similarly, researchers cannot randomly assign people or nations to different political cultures, economic circumstances, or global positions. Although experiments on these topics are likely to be limited to natural experiments and rare circumstances, the discipline has not yet tested the limits of randomized experiments. The causes of economic development, democracy, socialism, religious fundamentalism, or revolution may be too complex to be reduced to specific, and measurable, causal

relationships. On the other hand, studies of basic aspects of each phenomenon would be instructive.

When it comes to narrower research questions of interest to political candidates and organizations, it is important to note that “experimentation is possible whenever decision makers face constrained resources and are indifferent between competing ways of allocating them” (Green & Gerber, 2002, p. 821). Unable to reach every voter or donor, organizations could call as many as possible using a list of names ordered using random assignment. Those they do not have time to call become the control group. They can continue their work as usual while allowing a researcher to assess the effectiveness of their efforts in ways that help future campaigns.

Policy Implications

Political scientists can work with any policymaker who has discretionary resources and an interest in causal relationships. Any policy that will be phased in can be phased in using random assignment to create treatment and control groups of individuals, blocks, or cities. The ability to test the effectiveness of specific social interventions has major advantages for policymakers. Just as the Food and Drug Administration relies on randomized experiments (drug trials) to test the safety of pharmaceutical drugs before approving them for consumer use, public policy makers can look to randomized field experiments to provide similar tests of the effectiveness of various social or legal interventions. Researchers could use experiments to assess the effectiveness of efforts to reduce prejudice, reduce crime, increase conviction rates, rehabilitate convicts, promote recycling, recruit civil servants, recruit military personnel, reduce health care costs, improve health care quality, promote political participation, and decrease dependence on social services. Education, health care, criminal justice, welfare, national security, and the environment are a few of the policy areas that might benefit from additional scientific experimentation. Experiments can promote evidence-based public policy decisions. Experiments can also help elected officials determine the best way to educate ordinary citizens about public policy and to get them more involved in the policy-making process.

Future Directions

For most of the 20th century, political scientists rejected the notion that politics could be studied experimentally (Lowell, 1910). Experiments remained rare, until an increased interest in causality, new computer technology, and innovative scholars pushed the experimentation forward (Druckman, Green, Kuklinski, & Lupia, 2006). In the future, political scientists will increasingly move laboratory

experiments from the college classroom to more natural settings to include a wider range of subjects. They will also increasingly test laboratory findings out in the field. For example, the work on negative campaigning has moved into the field, with randomly selected voters or zip codes receiving negative campaign mail (e.g., Niven, 2006) or negative radio spots (Green & Vavreck, 2008). Experimental work will also become increasingly sophisticated. For example, studies of voter mobilization have gone beyond testing the relative impact of leaflets, door knocks, and phone calls to testing the effectiveness of different messages, messengers, and timing on different kinds of voters (see Green & Gerber, 2005, for a collection of research by scholars working in this field). New experimental studies are also beginning to look at more complicated social phenomena, including the importance of social pressure (Gerber, Green, & Larimer, 2008) and (online and offline) social networks in shaping voter behavior. To increase the realism and relevance of their work, scholars are also doing more to work with real political campaigns, organizations, and governmental agencies. The U.S. government has also begun funding more large-scale experiments in political science, a trend that may continue as budget constraints lead to an emphasis on evidence-based policy decisions. Finally, the increased prominence and visibility of experimentation will lead to more experimental research on topics outside the subfields of legislative politics, public opinion, and political participation.

Conclusion

Experiments allow political scientists to test the relationship between cause and effect. The experimental method is one way to learn more about the political world. By randomly assigning subjects to treatment and control groups, researchers can isolate the effect of a specific intervention on subjects' political attitudes, knowledge, or behavior. Randomized experiments can be conducted in the laboratory or in the field. Researchers also conduct so-called natural experiments by seeking out circumstances in which specific interventions affect populations selected as if at random. Although survey-based research continues to dominate the discipline, scholars are increasingly turning to experiments as a way to overcome the problem of self-reporting that can bias survey responses. Critics raise questions about the internal and external validity of experimental research. Proponents of the method argue that both concerns can be addressed through extension and replication. Building on previous research, political scientists are using experiments to answer increasingly complex questions about a wide variety of topics, including, but not limited to, the political effects of political advertising, racial attitudes, and voter mobilization campaigns. Political science has relied less heavily on experiments than have the related fields of psychology and economics. The end of the

20th century marked an increase in important laboratory-based experiments, while the 21st century witnessed a movement toward field experimentation. Experiments are now employed in work across the discipline and in interdisciplinary studies of politics.

Note

1. Gosnell assigned certain blocks to receive a letter urging adults to register to vote; Hartman assigned houses to receive leaflets offering either rational or emotional appeals for the Socialist party. These early studies were controlled field experiments, but treatments were not assigned on a purely random basis.

References and Further Readings

- Ansolabehere, S., & Iyengar, S. (1995). *Going negative: How attack ads shrink and polarize the electorate*. New York: Free Press.
- Bogus, C. T. (1992). The strong case for gun control. *American Prospect*, 3(10), 19–24.
- Brady, H. E., & McNulty, J. E. (2004, July). *The costs of voting: Evidence from a natural experiment*. Paper presented at the 2004 meeting of the Society for Political Methodology, Palo Alto, CA.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Dawes, R. M., Orbell, J. M., Simmons, R. T., & van de Kragt, A. J. C. (1986). Organizing groups for collective action. *American Political Science Review*, 80, 117–185.
- Druckman, J. M., Green, D. P., Kuklinski, J. H., & Lupia, A. (2006). The growth and development of experimental research political science. *American Political Science Review*, 100, 627–636.
- Dunning, T. (2008). Improving causal inference: Strengths and limitations of natural experiments. *Political Research Quarterly*, 61, 282–293.
- Eldersveld, S. J. (1956). Experimental propaganda techniques and voting behavior. *American Political Science Review*, 50, 154–165.
- Gerber, A. S., & Green, D. P. (2000). The effects of canvassing, direct mail, and telephone calls on voter turnout: A field experiment. *American Political Science Review*, 94, 653–663.
- Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large scale field experiment. *American Political Science Review*, 102, 33–48.
- Gilens, M. (1999). *Why Americans hate welfare: Race, media, and the politics of anti poverty policy*. Chicago: University of Chicago Press.
- Gosnell, H. F. (1927). *Getting out the vote: An experiment in the stimulation of voting*. Chicago: University of Chicago Press.
- Green, D. P., & Gerber, A. S. (2002). Reclaiming the experimental tradition in political science. In H. Milner & I. Katznelson (Eds.), *State of the discipline* (Vol. 3, pp. 805–832). New York: W. W. Norton.
- Green, D. P., & Gerber, A. S. (Eds.). (2004). Field experiments in the political sciences [Entire issue]. *American Behavioral Scientist*, 47(5), 485–728.
- Green, D. P., & Gerber, A. S. (Eds.). (2005). The science of voter mobilization [Special issue]. *Annals of the American Academy of Political and Social Science*, 601(1), 1–204.
- Green, D. P., & Gerber, A. S. (2008). *Get out the vote! How to increase voter turnout*. Washington, DC: Brookings Institution.
- Green, J. G., & Vavreck, L. (2008). *Advertising and information*. Unpublished manuscript.
- Hartman, G. W. (1936–1937). Field experiment on the comparative effectiveness of “emotional” and “rational” political leaflets in determining election results. *Journal of Abnormal Psychology*, 31, 99–114.
- Hurwitz, J., & Peffley, M. (Eds.). (1998). *Perception and prejudice: Race and politics in the United States*. New Haven, CT: Yale University Press.
- Iyengar, S., & Kinder, D. R. (1987). *News that matters: Television and American opinion*. Chicago: University of Chicago Press.
- Iyengar, S., Peters, M. D., & Kinder, D. R. (1982). Experimental demonstrations of the “not so minimal” consequences of television news programs. *American Political Science Review*, 76(4), 848–858.
- Kuklinski, J. H., Cobb, M. D., & Gilens, M. (1997). Racial attitudes and the “new South.” *Journal of Politics*, 59, 323–349.
- Kuklinski, J. H., Sniderman, P. M., Knight, K., Piazza, T., Tetlock, P. E., Lawrence, G. R., & Mellers, B. (1997). Racial prejudice and attitudes toward affirmative action. *American Journal of Political Science*, 41(2), 402–419.
- Lau, R. R., Sigelman, L., Heldman, C., & Babbitt, P. (1999). The effects of negative political advertisements: A meta analytic assessment. *American Political Science Review*, 93(4), 851–876.
- Lowell, A. L. (1910). The physiology of politics. *American Political Science Review*, 4(1), 1–15.
- McKelvey, R. D., & Ordeshook, P. C. (1990). Information and elections: Retrospective voting and rational expectations. In J. Ferejohn & J. Kuklinski (Eds.), *Information and democratic processes* (pp. 281–312). Urbana: University of Illinois Press.
- Nickerson, D. W. (2008). Is voting contagious? Evidence from two field experiments. *American Political Science Review*, 102, 49–57.
- Niven, D. (2006). A field experiment on the effects of negative campaign mail on voter turnout in a municipal election. *Political Research Quarterly*, 49(2), 203–210.
- Sapiro, V. (1981–1982). If U.S. Senator Baker were a woman: An experimental study of candidate images. *Political Psychology*, 3, 61–83.
- Sniderman, P. M., & Grob, D. B. (1996). Innovations in experimental design in attitude surveys. *Annual Review of Sociology*, 22, 377–399.
- Wattenberg, M. P., & Brians, C. L. (1996). *Negative campaign advertising: Demobilizer or mobilizer?* Los Angeles: Center for Research in Society and Politics. Retrieved June 5, 2009, from the eScholarship Repository: <http://repositories.cdlib.org/crisp/3>

63

FORMAL THEORY AND SPATIAL MODELING

DONALD M. GOOCH

Arkansas Tech University

In Greek mythology, Hercules is tasked with 12 impossible labors to regain honor and thus ascend to Mount Olympus as a god. The job of explaining formal theory and spatial theory in a brief, nontechnical essay is a labor of sufficient difficulty to make the search for the Golden Fleece pale in comparison. Given that this author has no transcendental gifts (though Hippolyta's belt may be around here somewhere), aspirations, or pretensions, this chapter eschews the impossible task of summarizing and explaining the entirety of formal and spatial theory. Instead, this chapter settles for the daunting yet mortal goal of a thorough yet concise introduction to some of the classical and contemporary works of the formal and spatial theories on politics and the concepts, definitions, and models on which those works rest (for a more complete treatment of formal theory and its contribution as a field of political inquiry, see Morton, 1999; Ordeshook, 1992; and Shepsle & Bonchek, 1997). Although Duncan Black (1958) may have understated the mathematical underpinnings of spatial theory as "simple arithmetic," it is as true today as it was then that the fundamental assumptions, intuitions, and predictions of formal and spatial theory can be grasped with a relatively basic foundation in mathematics such as algebra and geometry. Formal theorists employ a range of advanced mathematical concepts (i.e., integral calculus, matrix algebra, etc.) in their models. However, one does not need these to understand what formal theory is, what the

foundational principles of formal theory are, and the gamut of its predictions and conclusions regarding political institutions and behavior. To the extent possible without compromising the material, this chapter keeps the discussion broad and descriptive and thus accessible to the undergraduate reader.

What Is Formal Theory?

Formal theory is a field of inquiry that uses mathematical techniques to explicitly and precisely define theoretical concepts and the relationships between those concepts. Formal mathematics permits the systemizing of theory and thus permits precise deductions and synthesis and enhances the decidability of scientific propositions. Although the term *formal theory* is common parlance, it is also known as rational choice theory, public choice, positive political theory, political economy, the economic theory of politics, and a variety of other synonyms. Two of the primary branches of formal theory used frequently in political science are game theory and spatial theory. Game theory is concerned primarily with the strategic interaction of utility-maximizing actors in competitive and cooperative settings (see Chapter 64, "Game Theory," in this volume, for more information). Spatial theory, as the reader will see, examines the behavior of actors by representing beliefs,

positions, choices, and institutional contexts in terms of spatial distance (most commonly as points on the Cartesian plane). Although formal theory shares a foundation in mathematics and logic with quantitative methods, it is distinct from the traditional empirical inquiry of standard statistical methods. Formal theory seeks to explicitly define political concepts and derive logical implications from their interrelations. Traditional empirical methods assess the relationships between political concepts through direct statistical analysis. Nonformal theory underpins much of the empirical work in political science. A nonformal model suggests relationships between actors and institutions in the real world of politics using general and sometimes ambiguous terminology. This is not to suggest that traditional theorizing is necessarily bad or unrelated to that of formal theorizing. Indeed, there may be an underlying formal model in a generally stated theory of politics that, as Arrow (1963) notes, has not been expressed formally because of mathematical or linguistic limitations. A model is formalized when we use abstract and symbolic representations to explicitly state the assumptions of the model and from which can be derived equilibrium and comparative statics predictions (Binamore, 1990; Elster, 1986; Morton, 1999).

For example, a nonformal voting model might entail this proposition: Voters vote for viable candidates that share their beliefs and positions on issues. This seems to be a reasonable statement of the voting process. Yet there is a great deal of ambiguity in this statement. What does it mean for a candidate to be viable? To what extent must a candidate share the voter's beliefs and positions relative to the other candidates? How do voters assess candidate positions, and how do they relate them to their own beliefs? Furthermore, how important to the voter is the prospect that his or her vote will be decisive in the election? The nonformal model is silent or ambiguous on these questions. A formal model aims at explicitly defining the processes at work (in this case, the act of voting), the actors participating in the process (voters and candidates), and the gamut of alternative outcomes based on those choices (whether the citizen votes). Riker and Ordeshook (1968), operationalizing a spatial model of voting based on the classic median voter theorem developed by Downs (1957), give just such a formal model of voting.

According to Riker and Ordeshook (1968), an individual will decide to vote if and only if this equation holds true:

$$P * NCD + D \geq C,$$

where, for each voter, P = the probability that this person's vote will affect the outcome of the election, NCD = perceived net benefits of one candidate over another (net candidate differential), D = the individual's sense of civic duty, and C = costs associated with the act of voting (opportunity costs, driving time, gas, etc.).

This modeled cost-benefit analysis used by the voter hinges the act of voting on the difference among the candidates between the perceived spatial distance of the candidates' positions and that of the potential voter's own preferences, conditioned by the probability that the individual's vote will be decisive in the election. The relevance of the difference between candidates is dependent on the probability of a decisive vote. If the voter's vote is not decisive, then the candidate differential is essentially irrelevant to the outcome of the election from the perspective of the potential voter. This formal theory of voting uses mathematical notation to precisely relate the costs of voting to the benefits the voter receives from voting, and in so doing, it provides a nonobvious expected outcome that tells something interesting about the rational voter. As the probability of a decisive vote goes to zero, differences between the candidates on issues are eliminated from the calculus of the vote decision. This fact led scholars to predict that citizens wouldn't collect costly information on politics such as the policy positions of specific candidates or parties, a prediction confirmed by the significant political ignorance of voters found in surveys. Also, although many observers have decried the problem of low voter turnout in the United States, the Downsian and Riker-Ordeshook voting model suggests the real puzzle is that anyone votes at all.

Formal Theory, Quantitative Methods, and Empirical Inquiry

One way to think about the difference between formal theory and quantitative methods employed for empirical inquiry, given that both use the language of mathematics, is to put the distinction in terms of the scientific method. In the social sciences, the scientific method involves stating a research question of some importance to our understanding of social phenomena; developing theories as to the processes, actors, and interactions within the social context; using hypotheses derived from these theories for empirical testing; using techniques to test these hypotheses against real-world data; and publicly reporting the results of those tests. Formal theory in political science is oriented toward developing precise theories with specifically defined assumptions and the derivation of their implications (the so-called front end of scientific inquiry) while statistical methodology applies mathematical rigor to the testing of theories and hypotheses (the so-called back end of scientific inquiry). This dichotomy, although useful, isn't without its problems. Although it is true that the foci of formal theory and quantitative methods are distinct and have been historically pursued separately in political science, it is incorrect to assert that empiricists are unconcerned with precise theorizing and formal theorists are indifferent to empirical testing. Both formal theory and quantitative

methods are effective tools to employ in the study of political phenomena and, in combination, can produce significant contributions to the knowledge of politics (Barry, 1978).

The increasing role of formal theory in political science is not without its critics. The behavioral revolution in political science that drew the discipline away from informal normative theories and descriptive analysis inspired greater and greater attention to developing strong empirical measures of political phenomena. Many see formal theory as a distraction from real, and hence important, empirical analysis of politics. Albert Einstein once quipped that not everything that is important can be measured, and not everything that can be measured is important. To its critics, this sums up the problem with formal modeling. Its important theories cannot be measured, and what it measures is not important. Yet the formal theories of politics address many of the most important questions in politics: why citizens vote, how organized interests form, why democracies emerge, and why nations go to war. Furthermore, there is merit in assessing pure theory in its own right. Formal theory can reveal surprising and counterintuitive behavioral expectations and provide insights into political processes in important areas suffering from a scarcity of available data. Pure theory is often a precursor to the development of empirically testable measures. Though we lack—and may ever lack—a complete model of political behavior, both theory and empirics have a role in filling and bridging the gaps. Many of the first principles from which formal theories are derived are either undiscovered or only partially described and understood.

There is room in the discipline for both forms of inquiry. The ambition of political science is to provide pieces of the puzzles of politics with increasingly better developed and more rigorously tested models of behavior. Both formal theorists and empiricists have contributions to make to our understanding of politics. Where possible, it is best to precisely define both our theoretical expectations and our empirical tests of those expectations. It is difficult to test theories that lack precision or clear implications, and the ambiguity of these nonformal theories can result in conflicting and mutually exclusive tests. The practical usefulness of precise theories is lessened without ways to test them against reality. Theories that wander too far away from the real world of politics make the discipline less relevant to both policymakers and students of practical politics. The combination of the two approaches, where we use quantitative methodology to assess the predictions and comparative statics of formal models against empirical data (empirical implications of theoretical models, often referred to as EITM), is one of the more significant modern trends in political science and is an active field of inquiry in the discipline, coexisting alongside the more traditional behavioral and pure theoretic fields of inquiry.

Whether through the investigation of the empirical implications of formal models or the mind experiments of pure formal theory, formal models have much to contribute to the study of politics today. Why do two parties form

plurality electoral systems, how do two major parties in first-past-the-post electoral systems respond to the threat of entry by third parties, why do voters turn out, how many seats should a party seek to control in a legislature, can we get irrational aggregate social outcomes when society is composed of rational individuals, why and how do procedural rules in institutions such as legislatures matter, and why do individuals choose to join interest groups? These questions and more lend themselves to formal analysis (Downs, 1957; Olson, 1965; Ordeshook, 1992; Palfrey, 1989; Riker & Ordeshook, 1968).

Rational Choice and Its Foundational Assumptions

Formal theory is a deductive form of inquiry, deriving implications and relationships from established first principles. One such assumption is rational choice. Rationality, as it is generally conceived in formal modeling, is the assumption that individuals have a set of preferences and act intentionally, constrained and motivated by real-world contexts, consistent with those preferences. Individuals are instrumentally rational. Thus, a rational choice is not a person doing what you think they should do if you were in their shoes, such as staying home and studying (you) rather than going out to a party before the big test (them). Just because you would value getting a good grade on the exam more than having fun on a Friday night does not make the other person's decision to party irrational. It also does not mean having superhuman knowledge or being brilliant decision makers. Individuals order their complete preferences as they see fit, and they make choices aimed at getting the best possible outcome according to those preferences from their perspective, however imperfect that may be.

There are three important principles that undergird rational choice. The first is that of completeness or comparability. If one is to choose among possible alternatives, one has to know what all the alternatives are and be capable of comparing them to one another. The second is the mathematical principle of transitivity (if $A > B$ and $B > C$, then $A > C$). To make a rational choice, one has to be able to order preferences consistently. The transitive principle permits a rational choice because the interrelation between all of a person's choices makes sense. If an individual prefers pizza to waffles and waffles to apples, then it isn't rational to prefer apples to pizza. Third, rational choice models assume that individual actors are self-interested, in that they attempt to get the best outcome possible for themselves. This is also called *utility maximizing*, where *utility* is just a quantifying term for a benefit to the individual and *maximizing* means that the individual seeks to get the largest benefit possible. Now, this isn't to say that all potential choices meet the comparability and transitivity and maximizing requirements. Indeed, people's choices

are often uninformed, hurried, inconsistent, and emotional. However, behavior that is best modeled as intentional, self-interested, and maximizing across comparable and transitive preference orderings—true of many political choices and decisions—lends itself to rational choice analysis (Riker & Ordeshook, 1973).

This definition of rationality reveals another fundamental assumption of formal theory: methodological individualism. Most formal theories employ the individual as the fundamental unit of analysis.¹ An individual can have preferences and beliefs while groups, firms, and states cannot. Consider again the Riker–Ordeshook model of voting. It assumes methodological individualism. Note that the model defines the individual citizen's calculus in deciding whether to vote. The Riker–Ordeshook model is also a good example of an application of the rationality assumption. They presume that the voter will assess both the internal factors (preferences ordered across candidates) and external factors (the probability that the voter's vote will be decisive) in making a rational cost–benefit decision whether to vote.

The assumption of rationality is one of the more controversial aspects of formal theory. Many critics argue that human beings lack the capacity, the evolutionary development, and the necessary information to make rational decisions as conceived by formal models. Although this may be the case, it does not necessarily mean that rationality is a useless or even pernicious assumption in formal theory. Assumptions can be both unrealistic and useful in terms of either identifying puzzles (if *X* is the rational choice, why do most individuals choose *Y*?) or by reflecting an important aspect of decision making, even if it does not accurately represent many individual decision-making processes. All models are inaccurate to some degree. Models are crude, stylized approximations of the real world, intended to reflect some important aspect of politics rather than every aspect. Model airplanes fall short of the realism ideal in terms of material composition, scale, and functionality. They are made of plastic rather than steel and fiberglass. Key components of real airplanes are missing or misrepresented. Few models use jet fuel or have afterburners. Yet model airplanes have fundamentally contributed to our understanding of flight and the design of aircraft. Indeed, real flight would have been impossible without creative modelers like Leonardo da Vinci informing and inspiring practical developers such as the Wright brothers. The measure of a model of politics is not whether it perfectly approximates the real world, but rather its usefulness and parsimony in contributing to our understanding of politics (Morton, 1999; Ordeshook, 1992; Shepsle & Bonchek, 1997).

That said, there have been significant innovations that incorporate more realistic assumptions regarding individual behavior in formal models. One important modification is a move away from deterministic models to probabilistic models of choice. This chapter has noted that utility

maximization is a key component of rational choice models, where people assign utilities to the outcomes of choices, and the rational individual chooses the highest utility outcome. When an individual is highly confident that *X* action will lead to *Y* outcome, we say that individual is operating under the condition of certainty. However, in many contexts, an individual is uncertain as to what actions lead to what outcome. Rather, individuals make choices that may or may not lead to a particular outcome. In such instances, the individual is uncertain about what happens when he or she makes a particular choice. When the individual has a good sense of the likelihood of certain outcomes (say, a 75% chance of *Y* and a 25% chance of *Z*), we say that individual is operating under the condition of risk. When an individual has no idea what will happen or what is likely to happen, he or she is faced with the condition of uncertainty (Dawes, 1988). Under probabilistic conditions, it is particularly useful to assign numbers to outcomes. Formal theory defines these as utility. Quantifying the outcomes permits the incorporation of probabilistic decisions into models of behavior. Now, rather than choosing acts that necessarily produce a particular outcome, the individual chooses among lotteries where the utility from outcomes is conditioned on the probability of that outcome occurring. This expected utility theory is an important innovation in modeling behavior. An individual may value being crowned king of the world very highly, and thus, we would assign that outcome a high score in utility. However, given that the probability of that outcome approaches zero, the individual's expected utility from choosing the actions that might lead to ascension to world ruler are actually quite low. This is why a lottery jackpot in the millions of dollars doesn't require individuals to buy lottery tickets to maximize their utility. Indeed, buying a lottery ticket may yield a lower expected utility than using that money on a soda or a hamburger (Shepsle & Bonchek, 1997).

Although probabilistic models may be more realistic, they still assume that individuals are rational utility maximizers. Other theorists have relaxed the assumption of rationality itself. Although this chapter cannot give them full treatment, nonlinear expected utility, prospect theory, bounded rationality, learning, and evolutionary models use near or quasi-rational models of behavior (Morton, 1999). Bounded rationality incorporates decision makers with incomplete information, who have cognitive limitations and emotional responses that prevent or complicate utility maximizing based on the limited information they do have, and the complexity inherent to decision making (Jones, 2001; Simon, 1957). Herbert Simon, an early developer of boundedly rational models, says individuals “satisfice” rather than satisfy a preference ordering. An individual who satisfices doesn't consider all possible alternatives, but rather uses a heuristic to search among a limited number of choices at hand that need not contain the optimal available decision when all preferences are considered (Simon, 1957). Prospect theory is a psychological theory of

decision making where individuals evaluate losses and gains differently. Loss aversion, where individuals fear losses more than they value gains, is a concept from prospect theory, and it generated different predicted behavior than traditional expected utility theory (Kahneman & Tversky, 1979). Most formal models in political science, however, employ the traditional assumptions of rationality and methodological individualism in constructing their models of politics.

Solving and Testing Formal Models: Equilibria, Point Predictions, and Comparative Statics

After a formal model has been developed, the model is solved for predictions presented as theorems or results. The implications (or solution) of the model are deduced axiomatically from the assumptions and structure of the model itself. In most formal models relevant to political science, the researcher seeks to solve the model analytically. This involves the search for equilibria (stable outcomes). Where analytical solutions are not feasible or possible, obtaining numerical solutions through computer simulation is an option. If the formal model is game theoretic, then the interactions between the players are strategic. A common solution concept in this form of model is the Nash equilibrium, where each player's choice is optimal, given the choices of other players, and thus, no player has an incentive to change strategies within the game. Solving for the decision of the potential voter in the Riker–Ordeshook model of turnout in a two-candidate election yields the instrumental solution that the voter should vote for his or her preferred candidate only if the probability that the vote will be decisive exceeds twice the cost of voting.² It furthermore yields the nontrivial result that, even if the cost of voting is very small, the voter should vote only if the probability of his or her breaking a tie exceeds 2 in 1,000. Given an election with a large number of voters (e.g., a presidential election), the Riker–Ordeshook equation yields a prediction: no vote.

In evaluating a formal model empirically, by relating the model to data from the real world (e.g., election results, legislative votes, presidential vetoes, etc.), one can evaluate assumptions, predictions, and alternative models. The evaluation of assumptions is a validation of the relevance of the formal model. If an assumption of a model is violated frequently in the real world, then the scope of the applicability of that model is smaller. In evaluating predictions, one can look at point estimates. Point estimates are the values of the variables in the model when in equilibrium (models can predict one or multiple equilibria). Another method of evaluation is comparative statics, where changes in the endogenous variables of the model in equilibrium (dependent variable) vary with the values of an exogenous variable (independent variable). Finally, one

can assess models by looking at them in competition with other contrary formulations of the political phenomenon (Morton, 1999).

Public Choice: Democratic Theory, Institutions, and Voting Paradoxes

Between 1950 and 1965, the seminal and foundational works in formal and spatial theory were published. Among them are Arrow's (1963) *Social Choice and Individual Values*, Black's (1958) *The Theory of Committees and Elections*, Buchanan and Tullock's (1962) *The Calculus of Consent: Logical Foundations of Constitutional Democracy*, Riker's (1962) *The Theory of Political Coalitions*, Olson's (1965) *The Logic of Collective Action*, and Anthony Downs's (1957) *An Economic Theory of Democracy*. Each represents an important contribution to formal modeling and identifies important paradoxes or puzzles of logical political behavior, collective action, choice mechanisms, and democratic theory that continue to be the subject of innovative research today.

However, the study of choice mechanisms using mathematics actually began in the 18th century. Procedural problems in electoral systems led Condorcet and Borda to investigate the problem analytically. In the 1920s, Pareto would use mathematics to understand social phenomenon (Pareto, 1927; Shepsle & Bonchek, 1997). It is Pareto's efficiency concept that underlies Buchanan and Tullock's (1962) calculus of consent. These thinkers paved the way for the explosion of formal and spatial political theory in the 1960s. The works of these political economists formed the pillars on which modern public choice theory was built. Public choice theory (also called social choice) focuses on macroinstitutional factors and how the structure of government interacts and often conflicts with the aggregate preferences of the public. Olson's work on the collective action problems inherent to group formation, Downs's conclusions regarding the rational ignorance of voters, Riker's theory on minimum winning coalitions, and Buchanan and Tullock's treatise on the political organization of a free society are all significant contributions worthy of attention, but this chapter focuses on only one aspect of public choice theory as an illustration: voting behavior and electoral competition.

Condorcet was one of the first to apply mathematical modeling to the problem of making a collective decision among a group of individuals with preference diversity (not everyone wants the same thing). One of the common themes in public choice theory is the normative principle that choice mechanisms should reflect democratic values. One such mechanism intended to reflect a democratic choice is first-preference majority rule, where the top preference of the largest number of individuals is given effect as the decision on behalf of the collective. But there can be

multiple majorities, and which choice is made is dependent on the order in which the alternatives are presented, particularly in pairwise comparisons. If there is one choice that defeats all others in a pairwise vote, it is said to be a Condorcet winner. Condorcet identified a problem with majority rule when group preferences are intransitive. Although rational choice requires transitive individual preference orderings, this does not require that group preferences be transitive in the aggregate. When group preferences are intransitive (Condorcet's paradox), cycling can occur. When A defeats B and B defeats C and C defeats A, there is not one majority rule election that will produce the group or democratic preference, since no such preference exists (McLean & Urken, 1993). There is no stable outcome that a majoritarian procedure can produce under these conditions.

Condorcet examined just the special case of majority rule. Arrow (1963) looked at the problem more generally by making a few basic minimal assumptions about what a democratic process would entail: All preference orderings are possible including those with indifference, Pareto optimality, the independence of irrelevant alternatives, and nondictatorship. He asked whether it was possible to construct a method that would aggregate those preferences in such a way as to satisfy these conditions. Arrow's theorem asserts that there is no such possible choice mechanism. Rational democracy isn't merely impractical; it is impossible. What this means practically is that there is a trade-off between having a rational system that translates preferences to policy and the concentration of political power. Put another way, dictators are good for consistency. This is not to say that all social aggregation is unfair or irrational. Rather, there is no mechanism that can guarantee such an outcome in any given context. Arrow's result shows that democratic processes that yield socially coherent policy are a much more difficult proposition than had been thought.

The Spatial Theory of Voting: The Median Voter Theorem and Theoretical Modifications

As noted earlier, one of the major innovations of formal theory was the use of geometric space to represent political choices. Let's set out some of the basics of spatial theory from the outset. The standard spatial model depicts voting with Euclidean preferences in a one- or two-dimensional space. This means that political choice is represented as the choice of some point on a line or a two-dimensional space over which all the actors have preferences. Specifically, each actor, j , has an ideal point (top preference) on the line or space, prefers a point closer to this ideal point to one more distant from it, and is indifferent between two equally distant points. In the two-dimensional case, an actor's indifference curves are

concentric circles centered on his ideal point. Actor j 's preference set $P_j(x)$ is the set of points j prefers to x . Furthermore, in most spatial models, preference orderings are assumed to be single peaked (monotonic).

Although single-peaked preferences are helpful in producing social consensus in the absence of unanimity, Duncan Black (1958) demonstrates that they are also an important aspect of the spatial representations of politics. If one takes a group of individuals (voters in an election, legislators on a committee) who are considering a policy along one dimension (say, candidates in an ideological dimension or the amount of tax dollars to budget for defense spending), and their utility function is single peaked, then the outcome of this process is determined by the median voter—specifically, the committee member located at the center of the group on the relevant dimension determines the outcome. Geometrically speaking, Black's median voter theorem shows that the ideal point of the median voter has an empty win set. A win set $W(x)$ is the set of all points that beat (are collectively preferred to) x under a decision rule. If the ideal point of the median voter has an empty win set, then the median voter's preference commands a majority against all other possible points on the policy dimension.

The commanding stature of the median voter was first suggested by Harold Hotelling (1929) in predicting the geographic congregation of firms at one location, such as hot dog vendors on a street. Although Black (1958) studied committees, Downs (1957) adopted Hotelling's proximity model in his now famous median voter theorem in elections (MVT). The theorem states that the median voter in a single dimension cannot be defeated in a pairwise vote with full turnout and sincere voting (individuals vote according to their true preference ordering rather than trying to game the vote by strategically voting for a less-preferred alternative).

Both Downs's and Black's theorems suggest there is a centripetal force at work in politics. Downs predicted that party (or candidate) platforms would converge to the median voter's policy preference. It is called a *proximity* model because Downs assumed that voters used the spatial distance between themselves and candidates to determine whom they should vote for. Rational voters in this model vote for the candidate or party with a platform closest to their most preferred policies. Parties converge because that's where the votes are (Downs, 1957). But does this model accurately depict how parties behave in real elections?

There are a variety of complications that can prevent Downsian convergence of parties. Empirically speaking, there is evidence from a plethora of elections here and abroad where parties and candidates failed to converge to a single policy point or even a vector of policy points in a continuum of policies. Divergence appears to be the norm rather than the exception (Morton, 1993). Multiple dimensions are also a complication for the MVT (Riker, 1980).

The Downsian model assumes party competition occurs in a unidimensional space, but political competition can occur along multiple dimensions. After all, the cost of a policy is only one consideration when it comes to deciding how to authoritatively allocate resources. Fairness, effectiveness, efficiency, and other considerations can come in-to play. Ideology is one way to rate candidates, but what about affect (likeability), trust, and performance considerations? Nonpolicy attributes, or valence, may influence election outcomes (Groseclose, 2001).

Empirical and theoretical issues with the MVT have led scholars to develop more sophisticated formal models of party competition. Scholars have extended the spatial model of voting developed by Downs into multiple dimensions using a Euclidian model of utility functions (Enelow & Hinich, 1984; Hinich & Pollard, 1981). Plott's (1967) theorem suggests that a median voter result is possible in multiple dimensions but that the conditions for it are attenuated. It is dependent on radial symmetry among the alternatives. McKelvey's (1976, 1979) chaos theorem asserts that there is no majority-rule empty-win-set point in a multidimensional spatial setting other than Plott's special case. In other words, we can start at an arbitrary point in the space, and a majority can move us to any other point in the space. With no Condorcet winner, policy cycles endlessly (McKelvey). However, is this chaos ephemeral? Consider that policy cycling isn't frequent in legislatures. As Gordon Tullock (1981) famously queries, "Why all the stability?" It remains a point of contention, though Shepsle (1979) suggests that institutions impose policy stability through restrictive rules. Institutions may impose stability, but this merely changes the choice context (rules instead of policy). Ultimately, it is an open question requiring further theory and study.

Theories of candidate divergence suggest alternative specifications such as nonnormal voter distributions, directional logic, permitting third-party entry, valence, and turnout variance as bases for moving away from the stylized median voter model. Can the distribution of voters produce platform divergence on its own? Where multiple modes exist, candidate divergence may be optimal. The implications for polarization in the median voter model were anticipated by Downs (1957). He argued that the location of equilibria in an election would be dependent on the shape of the distribution of citizen preferences. Downs, on this at least, was wrong. The pure MVT with complete turnout and sincere voting predicts that the median voter is, in fact, a Condorcet winner: No position defeats the ideal point of the median voter in a pairwise vote, irrespective of distributional qualities (Black, 1958).

Other scholars have taken issue with the proximity calculus where voters choose the candidate or party that is closest to them. Rabinowitz (1989) argues that platform space cannot be represented in terms of an ordered and continuous set of policy alternatives. Rather than a policy continuum, the directional theory of voting suggests policy

alternatives are dichotomous, and thus, candidates are judged by their intensity and policy direction. Finally, the prospect of entry by a third party may cause parties to diverge from the median to discourage a third-party challenge on their extremes in a polarized electorate (Fiorina, 1999; Palfrey, 1984). Hinich and Munger (1994) develop a theory of ideology that permits party divergence. Incorporating previous modifications to the MVT, such as incomplete information and uncertainty in voter policy locations and candidate locations, Hinich and Munger argue that the creation and maintenance of an ideology by parties is a necessary component of political competition. In a political environment where Republicans have become much more consistently and strongly conservative (and likewise, Democrats and liberals have become more liberal), vote-seeking parties rationally diverge to create a credible ideology that they can sell to their constituents. Establishing an ideological flag at one of the poles in a bimodal distribution can account for platform divergence.

The MVT has received a lot of attention in political science and, as the reader has seen, a great deal of criticism. Although the MVT point predictions on turnout and convergence have been falsified, the comparative statics of the model have been validated in observed elections. The fact that most parties in most elections do not converge to a single point on the policy dimension is not a failure of the MVT as a model of politics, let alone a failure of formal theory. There is a centripetal force drawing parties to the center in American politics, and Downs (1957) gives a parsimonious explanation of why that is. Furthermore, a plethora of theoretical modifications and empirical tests have been conducted using formal theory in extending and critiquing the MVT that has greatly advanced political scientists' understanding of party behavior in elections. This is a key point. The efficacy of formal modeling is not dependent on the success or failure of one model. Indeed, one can have competing formal models with polar opposition predictions about the same phenomenon. Formal theory is a deductive method of social scientific analysis. The analyses using formal theory stand or fall on their own merits.

Conclusion

The major paradigms of formal and spatial theory in social choice, voting, institutions, and political behavior have spawned decades of empirical and theoretical research as well as countless additional, alternative, and contrary models of political decision making. Many of the early formal models discussed here have been either falsified or significantly modified to account for empirical deficiencies. Riker's (1962) theory of minimum winning coalitions doesn't describe many legislative contexts, and the behavior of legislators often violates his theoretical expectations. Downs's (1957) turnout prediction has been falsified, and

his prediction of party convergence at the median has been serially violated in actual elections. As Plott (1967) observes, Downs's and Black's median voter theorems are problematic in multiple dimensions. The cycling and instability of social choice identified by McKelvey (1976, 1979) is not a consistent characteristic of government institutions, causing some political scientists to puzzle over the apparent stability in these institutions and their incorporated choice mechanisms.

These empirical failures and the ad hoc modifications aimed at rescuing them have led some scholars to suggest that political behavior is inherently irrational or, at minimum, that there is a poverty of realistic and empirically supported rational choice models, rendering them to be of little use or relevance (Green & Shapiro, 1994). This is a mistake. It ignores important empirical validations of formal models, fetishizes point predictions over comparative statics, and sets up a straw man of rational choice theory when there is not one but rather a multitude of formal, spatial, and rational choice theories. The failure of one or more formal models does not prove that formal theory has little utility in empirical investigations of politics. Those failures spur puzzle solving, the development of better and alternative models, and the exploration of new and innovative empirical tests of model predictions.

One can see this in the variety of formal models, evidence, and arguments directly responding to the Downs–Hotelling proximity model of party platform convergence. The Riker–Ordeshook model of turnout that was considered at the beginning of the chapter modified the traditional Downsian turnout model by incorporating a new variable: a psychic benefit from participation. Political scientists have learned quite a bit from the so-called failure of the Downsian turnout and proximity models. They now know that instrumental calculation (the cost of voting combined with the probability of affecting the outcome) is insufficient to spur a voter to participate. Rather, the experiential benefit characterized as a psychic civic-duty benefit by Riker and Ordeshook (1973) is the decisive consideration. Formal modelers have incorporated alienation, abstention, and other modifications to account for positive turnout in elections. Thus, an apparent formal model failure has actually yielded numerous and significant contributions to our understanding of voting behavior. These and many other formal treatments of politics are real and important contributions to social scientific knowledge. For political scientists, formal and spatial models are essential tools for understanding and predicting political behavior and phenomena.

Notes

1. More recent innovations have moved away from the atomistic individual assumption in adopting institution and context specific models.

2. See Shepsle and Bonchek (1997), pp. 251–259, for a detailed solution of the Riker Ordeshook turnout model in a two candidate election.

References and Further Readings

- Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New Haven, CT: Yale University Press.
- Barry, B. (1978). *Sociologists, economists and democracy*. Chicago: University of Chicago Press.
- Binamore, K. (1990). *Essays on the foundation of game theory*. Oxford, UK: Blackwell.
- Black, D. (1958). *The theory of committees and elections*. Cambridge, UK: Cambridge University Press.
- Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent: Logical foundations of constitutional democracy*. Ann Arbor: University of Michigan Press.
- Dawes, R. M. (1988). *Rational choice in an uncertain world*. Orlando, FL: Harcourt Brace.
- Downs, A. (1957). *An economic theory of democracy*. New York: Harper & Row.
- Elster, J. (1986). *Rational choice*. New York: New York University Press.
- Enelow, J. M., & Hinich, M. J. (1984). *The spatial theory of voting: An introduction*. Cambridge, UK: Cambridge University Press.
- Fiorina, M. P. (1999, October). *Whatever happened to the median voter?* Paper presented at the Massachusetts Institute of Technology Conference on Parties and Congress, Cambridge.
- Green, D. P., & Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science*. New Haven, CT: Yale University Press.
- Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science*, 45(4), 862–886.
- Hinich, M. J., & Munger, M. C. (1994). *Ideology and the theory of political choice*. Ann Arbor: University of Michigan Press.
- Hinich, M. J., & Pollard, W. (1981). A new approach to the spatial theory of elections. *American Journal of Political Science*, 25, 323–341.
- Hotelling, H. (1929). Stability in competition. *Economic Journal*, 39, 41–57.
- Jones, B. D. (2001). *Politics and the architecture of choice: Bounded rationality and governance*. Chicago: University of Chicago Press.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- McKelvey, R. D. (1976). Intransitivities in multidimensional voting models and some implications for agenda control. *Journal of Economic Theory*, 12, 472–482.
- McKelvey, R. D. (1979). General conditions for global intransitivities in formal voting models. *Econometrica*, 47(5), 1085–1112.
- McLean, I., & Urken, A. B. (1993). *Classics of social choice*. Ann Arbor: University of Michigan Press.
- Morton, R. B. (1993). Incomplete information and ideological explanations of platform divergence. *American Political Science Review*, 87, 382–392.

- Morton, R. B. (1999). *Methods & models: A guide to the empirical analysis of formal models in political science*. Cambridge, UK: Cambridge University Press.
- Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Ordeshook, P. C. (1992). *A political theory primer*. New York: Routledge.
- Palfrey, T. R. (1984). Spatial equilibrium with entry. *Review of Economic Studies*, 51, 139–156.
- Palfrey, T. R. (1989). A mathematical proof of Duverger's law. In P. C. Ordeshook (Ed.), *Models of strategic choice in politics* (pp. 69–92). Ann Arbor: University of Michigan Press.
- Pareto, V. (1927). *Manuel d'économie politique* [Manual of political economics]. Paris: Marcel Giard.
- Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Review*, 57, 787–806.
- Rabinowitz, G. (1989). A directional theory of issue voting. *American Political Science Review*, 83(1), 93–121.
- Riker, W. H. (1962). *The theory of political coalitions*. New Haven, CT: Yale University Press.
- Riker, W. H. (1980). Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review*, 74, 432–446.
- Riker, W. H., & Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review*, 62, 25–42.
- Riker, W. H., & Ordeshook, P. C. (1973). *An introduction to positive political theory*. Engelwood Cliffs, NJ: Prentice Hall.
- Shepsle, K. A. (1979). Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science*, 32, 27–59.
- Shepsle, K. A., & Bonchek, M. S. (1997). *Analyzing politics: Rationality, behavior, and institutions*. New York: W. W. Norton.
- Simon, H. (1957). *Models of man: Social and rational; mathematical essays on rational human behavior in a social setting* (3rd ed.). New York: Free Press.
- Tullock, G. (1981). Why so much stability? *Public Choice*, 37(2), 189–204.

64

GAME THEORY

BYUNGWON WOO

The Ohio State University

Game theory is a branch of applied mathematics that is used to model multiactor interdependent decision making. Game theory is widely used in many social science disciplines, including political science, economics, sociology, and anthropology, where researchers are interested in outcomes when at least two actors interact with certain purposes.

Game theory is a method of modeling. A usual game theoretic model specifies some essential aspects of a situation of interest and tries to make logical inferences about ensuing outcomes given the initial setup. There can be a simple election model, for instance, where there are two candidates who want to win the election and n voters who want to elect the candidate who is going to make policies that are beneficial for the voters. Two candidates announce their respective policy platforms, and voters vote. Whoever gets the majority of votes wins and makes policies. Given the initial setting, the solution to the game provides logically deduced inferences about outcomes of interest, such as who can win under which conditions and which policies should follow.

In modeling a situation, a game theoretic model captures only essentials and inevitably leaves out unnecessary details. Thus, a game theoretic model does not and cannot perfectly reflect the reality. Thus, a game theoretic model may seem too abstract. Indeed, one of the common criticisms of game theory is that game theoretic models are too unrealistic. Yet abstraction is common for any kind of

modeling. For instance, a model of an airplane or a model of an automobile usually does not feature every nut and bolt of an actual airplane or automobile. Instead, the models would probably have a cockpits and wings for airplanes and doors, tires, and wheels for cars, yet they would probably not have emergency oxygen masks, all cockpit buttons, and smoke detectors in the restrooms of airplane models and cup holders and detailed electrical lines connecting batteries to different parts of car models. Yet these models may still be useful for certain purposes. Similarly, a game theoretic model capturing the relationship between Congress and bureaucrats may feature only two actors even though Congress and bureaucrats are not unitary actors but rather composed of groups of individuals in reality. Congress and bureaucrats in the model would also be assumed to have a few primary goals, such as reelection and budget maximization even though there are many other potential motivations for each actor. Yet as is the case for the automobile and airplane models, the model can still be proven to be useful to study the relationship between Congress and bureaucrats.

There is no golden rule as to how abstract or realistic a model should be. In addition, there are a number of ways to model a situation by emphasizing certain aspects of the situation at the expense of other aspects being bracketed. Thus, it is hardly possible to tell if a model in itself is either right or wrong. Instead, a model can be judged by how useful and applicable it is to a modeled situation. Generally

speaking, a researcher can make a model resemble the reality more faithfully with many details, but in doing so, the researcher has to face the usual trade-off between details and generalizability. That is, a detailed model may capture a particular situation more accurately but not be generalizable beyond the particular situation. A more abstract model, in contrast, may be more general and applicable to a broader set of situations but may seem too unrealistic to approximate a particular case. In addition, each addition of details would make the model more complicated and make the math difficult or even analytically intractable. Ultimately, the decision for the initial setup and how complicated a model should be depends on the researcher's purpose.

Game theory is often called a method. But it is important to acknowledge the difference between methods of modeling like game theory and methods of empirical testing, which usually refers to research methods. Essentially, game theory is a method of theorizing or modeling, as opposed to a method of empirical testing such as regression analysis, factor analysis, longitudinal analysis, and other qualitative and quantitative research methods introduced in other chapters in this volume. The critical difference is that a method of theorizing is used to generate hypotheses or expected relations between important and interesting factors under research investigation while a method of empirical testing examines if and how well the generated hypotheses match real data. Thus, the two methods are complementary. Often, scholars use a method of theorizing such as game theory in the theoretical section of a research study to generate hypotheses and then use an appropriate research method to test the hypotheses in an empirical section in a single study. For instance, a game theoretic model in a study may generate the hypothesis that states the following: As trade between Country A and Country B increases, the probability of an interstate conflict between A and B increases. Then an appropriate empirical method such as binomial logit or probit analysis in the study can be installed to see if the stated hypothesis holds in reality with actual trade and conflict data sets.

Game theory is used to capture multiactor interdependent decision making processes. Naturally, then, there should be more than one actor making decisions in the models. This differentiates game theory from decision theory, which models a single-actor decision-making process. Also, the ensuing outcome (and thus the ensuing payoff) of the multiactor decision making should be interdependent in a game theoretic model. In other words, the final outcome needs to be jointly determined by actors involved in the model. The rock-paper-scissors game is a good example. More than one player is needed to play this game, and the outcome is jointly determined by the decisions of both actors. Player A can take one of three actions: rock, paper, or scissors. Likewise, Player B can take one of the same three actions. And the outcomes, A wins and B loses, the

game is tied, or A loses and B wins, are jointly determined by the actions that Player A and Player B take. By assuming that both players want to win, one can model the rock-paper-scissors game, solve it, and make predictions about each player's optimal actions in the game.

Since many social scientific research questions are about outcomes that result when multiple actors interact, game theory can be very useful in making inferences about potential outcomes in multiactor decision-making situations. Indeed, game theory has become increasingly popular in many social science disciplines for the past half century or so. In economics, where the use of game theory had been accepted earlier than in political science, game theory has been applied to model interactions between different sets of economic actors. For instance, economists have used game theory to model behaviors of competing firms, wage bargaining between a labor union and management, behaviors of producers and consumers, and competition among bidders at an auction. In political science, scholars have applied game theory to model behaviors of competing candidates in an election; the interaction between candidates and voters; the interactions between a bureaucratic agency and Congress; the interactions between the executive and the legislative branches in American politics; behaviors of states involved in interstate militarized disputes; behaviors of states in trade disputes; alliance behaviors; the role of mediators in conflicts; negotiations among states over design of international organizations in international politics; negotiations between parties to form, continue, and dissolve a coalition government; intrastate conflicts between factions in a country; and interactions between a government and an opposition in comparative politics.

This chapter introduces the basics of game theory and reviews the use of game theory in political science. In the next section, a few basic components of game theory and important terminology are introduced. Then a few representative examples of the use of game theory in various political science contexts are discussed. The discussion focuses particularly on three representative examples drawn from each of three subfields of political science: American, comparative, and international politics. These examples are among the most well known and widely cited and have made major contributions to the understanding of political phenomena. In the concluding section, a recap of the chapter is provided, and the ongoing effort of moving game theoretic models forward and the future of game theory are briefly discussed.

Definitions and Basic Terminology

Game, Players, and Preferences

A game refers to a strategic situation that involves at least two rational individuals called players. A rational

player is one who engages in goal-directed behavior—one who has well-defined goals, such as vote maximization or profit maximization, can order her or his preferences over alternative outcomes given a set of alternatives, and chooses the best alternative(s) for the realization of the given goals. For instance, when we model a Congressional election, we would probably assume that any serious candidates want to maximize their vote shares and that winning the election is the candidates' goal. Then each would have a set of alternatives that he or she needs to make a choice over, such as where to spend time and energy during the campaign, given district A, B, C, and D. A rational player then would choose the best alternative that would allow him or her to increase the vote share the most. This does not necessarily mean that a rational player is greedy or pursues only materialistic benefits. Contrary to a popular misunderstanding, game theory is agnostic about the origin of goals; hence, goals for players in a game may well be altruistic or emotional.

Players are assumed to have clear goals when they are involved in a strategic situation. Given the goal, they are able to arrange their preference ordering over every possible outcome. For instance, players involved in a rock-paper-scissors game are assumed to have a goal of winning. Then their preference ordering would be presumably to win over to tie and to tie over to lose (hence, win over lose).¹ Usually, preferences are translated by some utility functions that assign payoffs (real numbers) to each outcome when outcomes are determined by the combination of actions by all players. Payoffs for the outcomes are assigned such that the preference relations are maintained, given a set of outcomes and actors' preference relations among the outcomes. Following the rock-paper-scissors example, the utility function may assign a real number for each outcome so that each player gets a payoff of 5 for winning, 0 for tying, and -5 for losing. Since 5 is greater than 0 and 0 is greater than -5 , the preference relationship still holds.

We play many games in our everyday lives; games are often being played when people interact. For instance, if an individual drives a car in a busy street, that individual plays a game with the drivers of the other cars. Most drivers have clear goals: They want to spend the least amount of time on a road without being involved in a car accident, and each driver's decisions to change lanes, to stop at a light, and to choose one road over another all affect others' driving time on a road. Similarly, when an individual makes a bid for a pair of concert tickets on eBay, he or she is playing a game with other bidders. The individual would want to win an auction with a minimum bid, others would do the same, and one bidder's bid affects others' willingness and the price of bidding since each bid would affect the others.

Many political situations in real life can be thought of as games. The decisions of the Soviet Union and the United States about developing, stockpiling, and locating

nuclear weapons during the cold war era can be modeled as a game between the two superpowers. Similarly, the decision making by Nikita Khrushchev and the Soviet government during the Cuban missile crisis to build and remove a missile base on Cuban soil and the decision-making process by John F. Kennedy and the United States government to respond to the attempted construction of the missile base can be modeled as a game between two players pursuing their own security interests and interacting with various policy alternatives.

Game Representation and Solution Concepts

Formally speaking, a game consists of (a) a set of players, (b) a set of actions (or combinations of actions called *strategies*) for each player, and (c) preferences over the set of action (or strategy) profiles for each player. And a game is usually represented in one of two ways: normal form or extensive form.

Normal Form

The normal-form representation of a game specifies the players, the actions or strategies—the combinations of actions—for each player, and the payoff received by each player in a matrix. This is useful to represent situations where players make a strategic decision without knowing other players' decisions. A well-known example, the prisoner's dilemma game, is presented in the normal form in Table 64.1.

	Prisoner 2 Keeps Quiet	Prisoner 2 Confesses
Prisoner 1 Keeps Quiet	1, 1	9, 0
Prisoner 1 Confesses	0, 9	6, 6

Table 64.1 Prisoner's Dilemma

The background story of the prisoner's dilemma game is as follows. Two prisoners have been caught and are being interrogated by the police. The crime that they have committed and been caught for is relatively minor, but they have also committed a more serious crime in the past, and the police interrogate the prisoners to prosecute them for the serious crime as well. The prisoners are interrogated separately, without a way to communicate or collude with each other. The deal that the police propose to each prisoner is that if both prisoners remain quiet for the serious crime, both prisoners will serve only 1 year each in prison for the light crime without being convicted for the serious crime; if one prisoner remains quiet for the serious crime but the

other one confesses the serious crime, the one who confesses is set free while the one who remains silent serves a 9-year prison term; and if both prisoners confess, both are prosecuted for the serious crime and serve 6 years each in prison.

Here, two players, Prisoner 1 and Prisoner 2, are playing the game. They need to decide whether to keep quiet or confess, without being sure about the other player's decision. In the normal-form representation, players are usually listed on the top and the left side of the payoff matrix. Available actions are defined in columns and rows, and respective payoffs are listed in each cell where actions by players intersect. For instance, if Prisoner 1 remains quiet while Prisoner 2 confesses, Prisoner 1's payoff is -9 and Prisoner 2's payoff is 0.

Given the setup, then, what is the best strategy available for each player? There are possibly many solution criteria, but one intuitive strategy would be a simple procedure of elimination of dominated strategies. The simple procedure of elimination of dominated strategies relies on the reasoning that a rational player should not choose a strategy if there exists an alternative strategy that raises his or her payoffs against all possible strategies of his or her opponent (McCarty & Meirowitz, 2007). If we apply the procedure to the prisoner's dilemma, we can indeed obtain a solution. Suppose that Prisoner 1 believes that Prisoner 2 will confess. If that is the case, then it is better for Prisoner 1 to confess as well, because Prisoner 1 gets -6 by confessing as opposed to getting -9 by remaining silent. Similarly, suppose that Prisoner 1 believes that Prisoner 2 will remain silent. If that is the case, then it is better for Prisoner 1 to confess, because Prisoner 1 can get 0 by confessing as opposed to getting -1 by remaining silent. Thus, remaining silent is always dominated by confessing. Thus, the strategy of remaining silent can be safely eliminated from Prisoner 1's choice set. Since the game is symmetric and both players reason the same way, Prisoner 2 can also eliminate the option of remaining quiet. As a result, the only combination that remains possible is confession for both players. In general, if we repeat the procedure of elimination of dominated strategies, we may get a solution or at least tighten our predictions by eliminating a few strategies.

A more formal solution concept that is commonly used is the Nash equilibrium. Formally, Nash equilibrium is defined as "an action profile a^* with the property that no player i can do better by choosing an action different from a_i^* , given that every other player j adheres to a_j^* " (Osborne, 2004, p. 22).² In the prisoner's dilemma game, the action profile of the confess-confess combination is one (and only one) Nash equilibrium since there is no incentive to deviate from the confess-confess action profile for either player given that the other player sticks with confessing. That is, one becomes worse off only by deviating because if one player decides to deviate and remain silent, he or she will receive -9 as opposed to -6 in the confess-confess

case. In all other action profiles, however, each player can be better off by deviating from the profiles, given the other player sticking with the action; hence, these action profiles are not Nash equilibria.

Extensive Form

An extensive-form representation of a game involves the same set of elements in a normal-form representation—a set of players, a set of actions or strategies, and preferences for each player over a set of possible outcomes—with the addition of sequences. Thus, an extensive-form representation of a game is especially useful when one needs to explicitly take into account the sequence of actions by players. A game tree is commonly used to graphically represent a game (see Figure 64.1).

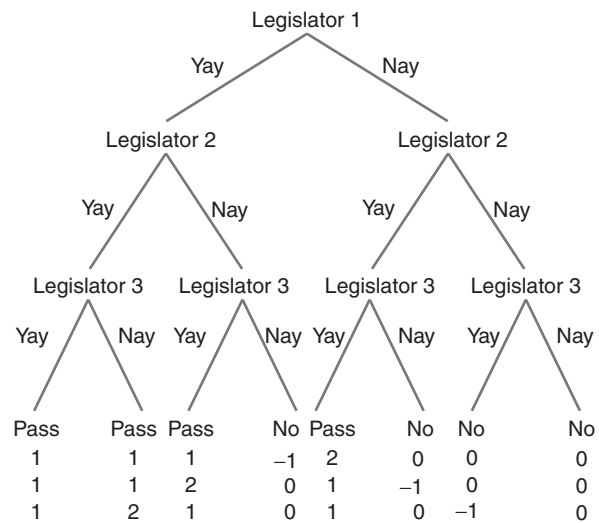


Figure 64.1 A Game Tree

The game is a simple legislation game among three legislators. The specific context of the game is as follows. Three legislators vote for the legislation to raise the pay for legislators. All legislators are assumed to want to get a pay raise, but they also do not want to be seen as pursuing their own self-interests by their respective voters. Thus, although all want to get a pay raise, each wants the bill to pass with others' votes but not with his or her own vote. Assuming a simple majority rule, this would imply the preference ordering of (a) bill passage with voting no, (b) bill passage with voting yes, (c) bill nonpassage with voting no, and (d) bill nonpassage with voting yes, the least favorite for all legislators. Now let's further assume the utility function that assigns a value of 2 for bill passage and -1 for voting yes, which is consistent with the description above. As a result, the first option gives 2, the second option gives $2 - 1 = 1$, the third option gives 0, and the fourth option gives -1.

The players, actions, and payoffs are all represented in the game tree. Players, Legislators 1, 2, and 3, are listed at each node. Although there seem to be multiple Legislator 2s and 3s, it is just the way it is represented, and there is actually only one Legislator 2 and one Legislator 3 playing at every possible node. Actions, *Yay* and *Nay*, are listed above the branches of the game tree. Outcomes, *Pass* or *No*, are followed from the combinations of actions taken by all players, and the payoffs are listed such that the top one is the payoff for the first legislator, the middle one is for the second legislator, and the bottom one is for the third legislator.

Again, there are potentially many solution criteria, yet the most intuitive one is to move backward and find solutions. This procedure is called *backward induction*, since it starts at the bottom and moves backward to find a solution. The rationale for the procedure is that when a player has to make a decision, the player will predict actions that the players in the game will subsequently take if they all act rationally, and the player will choose the action that maximizes the payoff for him or her.

Applying the backward induction to the game drawn in Figure 64.1 yields an equilibrium. Starting at the bottom nodes where Legislator 3 has to make a decision, Legislator 3 will choose “Nay” at the left node since it gives 2 compared with 1, choose “Yay” at the second-from-the-left node since it gives 1 compared with 0, choose “Yay” at the second-from-the-right node since it gives 1 compared with 0, and choose “Nay” at the far-right node since it gives 0 instead of -1. Moving up the tree, now Legislator 2 makes the decision knowing how the rational Legislator 3 would make a decision. Since the legislator knows that Legislator 3 will choose “Nay” at the left node and “Yay” at the second-from-the-left node, if Legislator 2 votes “Yay,” Legislator 2 gets 1; if Legislator 2 votes “Nay,” Legislator 2 gets 2. Thus, at the left node, the rational choice for Legislator 2 is to vote “Nay.” Similarly, knowing that Legislator 3 will vote “Yay” at the second-from-the-right node and “Nay” at the far-right node, it is in Legislator 2’s best interest at the right node to vote “Yay,” receiving 1 instead of 0. Finally, moving another branch up, Legislator 1 has to make a decision. Now, Legislator 1 knows that Legislator 2 would choose “Nay” and Legislator 3 would choose “Yay” after Legislator 2 votes “Nay” if Legislator 1 votes “Yay,” and Legislator 1 knows that Legislator 2 chooses “Yay” and Legislator 3 chooses “Yay” after Legislator 2’s “Yay” vote if Legislator 1 votes “Nay,” so it is in Legislator 1’s best interest to vote “Nay” and receive 2 instead of voting “Yay” and receiving 1. Thus, the equilibrium for the simple voting game is that Legislator 1 votes “Nay” knowing that the rest of the legislators will still vote “Yay” and pass the bill to raise their salaries, and the rest indeed vote “Yay” since they still prefer to vote “Yay” and pass the bill. Thus, the first voter, Legislator 1, enjoys a clear advantage, often referred to as the first-mover advantage.

Formally, the solution concept for the extensive-form game with complete information is a subgame perfect equilibrium. A subgame is defined as follows: “For any nonterminal history, h , where history is defined as a sequence of actions taken thus far, the subgame following h is the part of the game that remains after h has occurred” (Osborne, 2004, p. 164). For instance, after the history of Legislator 1’s playing “Yay” and Legislator 2’s playing “Yay,” the subgame following “Yay–Yay” history is the game in which Legislator 3 decides to vote “Yay” or “Nay.” Then a subgame perfect equilibrium is defined as “a strategy profile s^* with the property that in no subgame can any player i do better by choosing a strategy different from s_i^* , given that every other player j adheres to s_j^* ” (p. 165). In other words, a subgame perfect equilibrium is “a strategy profile that includes a Nash equilibrium in every subgame” (p. 166).

Extensions

There are many possible extensions to the games that are presented in this chapter. One particular extension concerns the information assumed in the game. In the extensive-form game, the informational assumption is that every player involved in the game knows everything about the game; in particular, players know each other’s payoff structure. This is called complete information. In contrast, a game of incomplete information, a very common extension of a complete-information game, deals with a situation where at least one player is uncertain about others’ payoff functions and thus tries to learn the other players’ types. A usual game of incomplete information posits that there is at least one player who can be one of two or more types, where each type corresponds to a different payoff function that the player might have. Then given the initial belief about the probability distribution over the types of the player, uninformed players update their beliefs about the probability distribution over the types of the player in question and make decisions based on their updated beliefs. Commonly, when an informed player moves first, the game is called a signaling game, and when an uninformed player moves first, the game is called a screening game.

Players are commonly assumed to follow the Bayesian rule when updating their beliefs about the probability distribution over the types of another player. A commonly administered solution concept is the perfect Bayesian Nash equilibrium. Essentially, it requires that a strategy-and-belief pair be consistent and mutually reinforcing.

Another common extension to the basic games is the repeated game. When players interact repeatedly, each player can condition his or her action on the other players’ prior actions. Thus, one may expect that different dynamics emerge when players engage in a strategic game repeatedly. For instance, when the prisoner’s dilemma game is played once, each player’s optimal strategy is to defect as shown previously.

Game Theory in Political Science

A Brief History

Game theory has become increasingly popular in political science since it was first introduced to the discipline in the 1950s and 1960s. An initial political science application of game theory was a group decision among a large number of actors or voters. From there, scholars have established the canonical theorems such as the median voter theorem and Arrow's general possibility theorem, also known as Arrow's impossibility theorem. These initial developments were modified and advanced to study the U.S. Congress and its committees' decision making. Along with the development of theories about group decision making, often referred to as *cooperative game theory* or *social choice theory*, the development of bargaining games and noncooperative game theory found more applications across the subfield areas in political science. There are numerous applications of noncooperative game theory. For example, in American politics, scholars study campaign strategies of candidates in an election, or they study how Congress delegates some authority to an independent bureaucratic agency and controls it by monitoring and punishing if necessary. In comparative politics, scholars study how parties bargain over coalition-government building and ending coalition governing. In international relations, scholars study why and under what conditions states go to war and at what times states initiate trade disputes through the WTO's dispute-settlement mechanism. As the number of studies using game theory to build theories increases, it is very common to find a few articles in any issue of the leading political science journals, such as the *American Political Science Review* or the *American Journal of Political Science*, that use game theory to explain political phenomena of interest.

Game Theory in American Politics

In American politics, game theoretic models have been used in various political contexts. Examples include models of agenda setting, legislative bargaining, collective goods and particularistic goods provisions, lobbying, congressional committees, parties and elections, and bureaucratic agencies and legislature. Among these numerous models, one of the most well-developed research strands is the model of legislative bargaining. Since the seminal piece by Baron and Ferejohn (1989), there has been steady progress made in the use of game theory in legislative bargaining, with additional assumptions making models of congressional bargaining more accurately reflect institutional features of the U.S. Congress.

The Baron and Ferejohn (1989) model starts with an observation that the distribution of pork barrel projects is focused on a few states and localities, but the taxes used to fund pork barrel projects are widely spread throughout

American legislative districts. To provide a logical answer to the question of why pork barrel projects are distributed in such way, Baron and Ferejohn present a legislative bargaining game.

The game features n members in the legislature, and each represents a legislative district. The task for the n -member legislature is to decide the distribution of benefits across all legislative districts. Each member has an equal probability to be recognized to make a proposal at the beginning of a session. The proposal may be amended depending on the amendment rule. The amendment rule can be either open or closed. Finally, the proposal is voted on and is passed if the majority casts yes votes. In the following, only the game under a closed rule is discussed.³

At the beginning of the game, a legislator is randomly chosen to make a proposal. A chosen legislator gets to propose a division of the benefits across all legislative districts, and the rest of the legislators get to vote yes or no for the proposal. When one makes a proposal, then it is brought to a vote, and following the simply majority rule, the proposal passes or fails. When the proposal passes, the legislature adjourns, the benefits are allocated as proposed, and the game ends. When the proposal fails to garner a majority of votes, another random draw is made, and another one out of n legislators is chosen to make a proposal. The game continues until a proposal is voted to pass. The utility for each legislator is defined such that each legislator enjoys the benefit allocated for his or her district since the legislator can bring the allocated money for the district, which increases the legislator's probability of reelection. The benefit is discounted by a discount term, meaning that a legislator prefers the benefit in the present legislative session over the same benefit in the next legislative session.

Solving the game involves the backward induction procedure. Baron and Ferejohn (1989) find that there are multiple Nash equilibria and multiple subgame perfect Nash equilibria in this game, and they suggest a refinement. The refined solution concept that they provide to generate unique predictions for the game is called *stationary subgame perfect Nash equilibrium*. Essentially, the refinement provides a restriction such that only the strategies that are time independent remain.

The stationary subgame perfect Nash equilibrium suggests that under the closed rule, only a minimum winning coalition is formed. More specifically, the equilibrium for the game features that a recognized member proposes to receive $1 - \delta(n - 1)/2n$ and offers δ/n to $(n - 1)/2$ other members selected at random, each member votes for any proposal in which at least δ/n is received, and the first proposal receives a majority vote, so the legislature completes its task in the first session. Thus, the randomly chosen proposer enjoys a huge advantage since he or she can bring approximately half of the entire benefits to his or her district. The chosen legislators, who vote yes, can bring some benefits to their districts. The remaining half of the

legislators do not receive any allocation of benefits and vote no but fail to stop the bill from passing.

Since the publication of the Baron and Ferejohn (1989) model, a number of studies have built on the initial model. For instance, Baron (1991) considers a similar model with taxation coming into the utility function of each legislator. More recently, Volden and Wiseman (2007) consider legislative bargaining over division between collective goods and particularistic goods, then the distribution of particularistic goods. Overall, these series of studies of legislative bargaining have contributed to our understanding of legislative bargaining and the importance of legislative rules and institutional settings.

Game Theory in International Relations

In international politics, game theory has been used in illustrating the logic of anarchy and explaining the importance of relative and absolute gains, causes of war, problems of credible commitments and signals, the role of mediators in international conflicts, and the role of international organizations in various contexts. One of the pioneers in international relations is Bruce Bueno de Mesquita (1981, 1985). In a series of articles and books published since the late 1970s, he has made major contributions in the study of war and peace by modeling the decision making of a rational foreign policy maker to initiate a military conflict with an expected utility framework. Since then, there have been many game theoretic studies of interstate conflicts, and one of the most cited articles is James Fearon's (1995) article on rational explanations for war.

Fearon (1995) questions why states go to a war even if waging a war is seemingly not an optimal choice for either state in most conflict cases. Capturing a war between states as a bargaining, he points out that agreeing on a division of a coveted good, such as a piece of territory, without fighting a war is often optimal for both states, rather than fighting a war and dividing the coveted good. For instance, if two states have a territorial dispute over a territory, π , and the two states can fight and divide π into p and $\pi - p$ for State A and State B after the war, it is better for both states to not fight a war, and divide π into p and $\pi - p$, than to fight and divide π , as a war is costly for both states. Fearon then provides several conditions under which states might go to war even if there is a potential bargaining space where both are better off without fighting a war.

The game theoretic model of war by Fearon (1995) features two players, State A and State B, who have preferences over a set of issues, such as a disputed territory, represented by the interval between 0 and 1. State A prefers issue resolutions closer to 1 while State B prefers resolutions closer to 0. Let p be the probability of State A winning a war if two states fight a war and $1 - p$ be the probability of State B winning the war. And let c for each state be the cost of a war, due to battle casualties, revenue

spent on a war, and destruction caused by a war, assuming c positive means that a war is costly for each state.

In its simplest form, the game flows as follows. At the beginning of the game, State A demands the division of 1, called x . State B receives the demand x from State A and chooses whether it wants to fight or back down. If State B chooses to fight, a war occurs and the payoffs are distributed for States A and B such that State A gets $p - c$ and State B gets $1 - p - c$. If State B chooses to back down, then a settlement occurs with State A getting x and State B getting $1 - x$.

With slight modifications to the model, Fearon (1995) shows that there are three possible explanations for war. First, rational leaders may choose to go to war because they cannot locate a mutually acceptable settlement because of private information about the probability of winning and incentives to misrepresent private information. Second, rational leaders may decide to wage a war because of a commitment problem. Even if both states can agree on the terms of settlement, the division of 1, there is no enforcement mechanism in international politics that prohibits the state that becomes stronger after the settlement to demand more in the future. Anticipating this, the state on the losing side may prefer fighting a war to agreeing on the settlement. Finally, states may find it difficult to find a peaceful settlement because the issues under contest are indivisible.

Since the seminal article by Fearon (1995), many studies have built on Fearon's model to study the dynamics of war. For instance, Alastair Smith (1998) offers a similar war model with each battle providing additional information about the probability of winning a war for both sides and both states making optimal decisions with updated beliefs about the probability of winning a war.

Game Theory in Comparative Politics

There are a large number of studies in comparative politics that use game theoretical models to capture democratic policy making, with particular focus on comparisons between different institutional structures used across democracies. These institutional features include electoral systems such as specific voting rules and electoral district sizes, party systems, and relations between the legislature and bureaucrats. Other applications of game theory in comparative political settings include democratization and market reforms, especially how an incumbent government and an opposition interact in these political and economic transitions. Geddes (1991) provides one such model. She develops a game theoretic model of bureaucratic reform in Latin American democracies where politicians interact in deciding whether to support administrative reform. She derives propositions that reforms are more likely to occur when patronage is evenly distributed among the strongest parties and when the electoral weight of the strongest parties remains stable. She then tests the predictions with brief case studies of five Latin American countries.

One of the more well-known game theoretic models in comparative politics is the article by David Austen-Smith and Jeffrey Banks (1988) published in the *American Political Science Review*. The article, titled "Elections, Coalitions, and Legislative Outcomes," brings together a model of election and a model of policy decision in a legislative setting and provides a comprehensive model of electoral and legislative behavior. As voters vote in anticipation of policy outcomes such as tax, education, and health policies and as policies are determined within a legislature where rational politicians condition their policy platforms on the prospect of electoral success, the two processes are clearly linked, and modeling both political processes simultaneously certainly advances a better understanding of elections and policy making in a legislature.

In their game, there are three parties competing for votes from n voters. At the onset of the game, the parties simultaneously declare their policy positions over a one-dimensional policy space. This one-dimensional policy space can conveniently be thought of as the left-right ideology spectrum. When the parties announce their respective positions, n voters cast their votes for parties. The electoral system is the proportional representation adopted in some European countries, and legislative seats are allocated to each party according to the proportion of votes that each party receives if a party receives more than s votes.⁴ In their model, Austen-Smith and Banks (1988) assume that every party receives at least s votes to reduce unnecessary complication of the model. When seats are allocated by the numbers of votes that each party receives, the parties try to form a government. The party with the largest number of seats first proposes a composition of the governing coalition, distribution of benefits among coalition members, and a policy to be implemented. On receiving the proposal, the members of the coalition either accept the proposal or reject the proposal. When the members of the coalition accept the coalition, the coalition government is constituted with the proposed policy subsequently implemented. If the members of the coalition reject the coalition proposal, then the party with the second-largest number of votes proposes a coalition, a policy, and a distribution of benefits. If the proposed package is accepted by the members of the coalition, the government coalition is formed, the benefits are distributed, and the policy is implemented accordingly. If the proposed package is not accepted, then the last party gets to make a proposal. If a coalition government is still not formed after the last party makes a proposal, then a caretaker government is implemented that makes equitable policy and benefit-distribution decisions.

Voters are assumed to be policy oriented, with their utility functions defined such that they prefer a policy closer to their own ideal policies. Parties are assumed to enjoy allocated benefits when they are included in a coalition

government, and parties want to minimize the distance between their electoral platforms and the implemented policy.

Solving the game is complicated but logically straightforward since one needs only to follow the backward induction methods presented previously. In the equilibrium, it will always be the case that the majority party, if one exists, forms a government by itself, and if the party with the highest number of seats does not enjoy a majority, then the parties with the highest and the lowest number of seats form the governing coalition. This is because the party with the highest number of seats finds it cheaper to offer a coalition to the party with the lowest number of seats.

Conclusion

This chapter provides a gentle introduction to game theory. As a tool for researchers to deduce logically consistent hypotheses, game theory has been widely used in many different social science contexts. Basic terms and elements of game theory and the most important solution concepts are introduced with some sample applications. Then three representative examples in political science are provided in the latter part of the chapter. One can see that game theoretic models can be used to study many interesting political phenomena, including legislative bargaining in American politics, decisions to go to war in international relations, and formation of coalition governments in comparative politics. For its relatively short history in political science, the influence of game theory on the ways in which researchers approach research questions has been substantial.

There are also ongoing innovations in game theoretic applications in political science that look very promising. One such innovation is to incorporate insights from psychological research in specifications of utilities for players and their ways of processing information. Another innovation is to allow players in a model to make various errors. Often referred to as *bounded rationality models*, these models are often made to allow better reflection of reality in a model.

Also innovative in political science is an effort to test theoretically generated insights empirically. Often dubbed *empirical implications of theoretical models*, there have been impressive attempts to bridge the gaps between theoretical propositions and empirical testing, including statistical and qualitative research methods. It is well acknowledged in political science that a multimethod approach to a given research question often yields a better result, and the ongoing effort to bring these theoretical and empirical research tools to study a research topic will certainly help us better understand and analyze complex political phenomena.

Notes

1. This is a property called *transitivity*. Generally speaking, if one prefers A over B and B over C, then it should follow that one prefers A over C to meet the transitivity condition. See Chapter 5, “Rationality and Rational Choice,” and Chapter 63, “Formal Theory and Spatial Modeling,” in this volume.

2. Action profile a^* is composed of the action by player i , denoted as a_i^* , and the action by player $j(s)$, denoted as a_j^* .

3. Under an open amendment rule, the game becomes a bit more complex. Essentially, when a proposal is made by a proposer, another random draw is made, and the chosen member of the legislature can either offer an amendment or suggest to move the proposal for voting. When the latter happens, a vote occurs. When the former happens, voting between the original proposal and the amended proposal occurs, and another member is randomly recognized to offer another amendment or move the previous proposal. The game continues until the majority of the n members vote yes on a proposal.

4. Many countries with a proportional representation system have the minimum vote threshold. Parties need to receive more votes than the threshold to be allocated legislative seats.

References and Further Readings

- Austen Smith, D., & Banks, J. (1988). Elections, coalitions, and legislative outcomes. *American Political Science Review*, 82, 405–422.
- Baron, D. P. (1991). Majoritarian incentives, pork barrel programs, and procedural control. *American Journal of Political Science*, 34, 57–90.
- Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. *American Political Science Review*, 83, 1181–1206.
- Bates, R. H., Greif, A., Levi, M., Rosenthal, J. L., & Weingast, B. R. (1998). *Analytic narratives*. Princeton, NJ: Princeton University Press.
- Bueno de Mesquita, B. (1981). *The war trap*. New Haven, CT: Yale University Press.
- Bueno de Mesquita, B. (1985). The war trap revisited: A revised expected utility model. *American Political Science Review*, 79, 156–177.
- Fearon, J. D. (1994). Domestic political audiences and the escalation of international disputes. *American Political Science Review*, 88, 577–592.
- Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, 49, 379–414.
- Geddes, B. (1991). A game theoretic model of reform in Latin American democracies. *American Political Science Review*, 85, 371–392.
- Gibbons, R. (1992). *Game theory for applied economists*. Princeton, NJ: Princeton University Press.
- Hinich, M. J., & Munger, M. C. (1997). *Analytical politics*. Cambridge, UK: Cambridge University Press.
- McCarty, N., & Meirowitz, A. (2007). *Political game theory*. Cambridge, UK: Cambridge University Press.
- Morrow, J. (1994). *Game theory for political scientists*. Princeton, NJ: Princeton University Press.
- Osborne, M. J. (2004). *An introduction to game theory*. New York: Oxford University Press.
- Powell, R. (1999). *In the shadow of power: States and strategies in international politics*. Princeton, NJ: Princeton University Press.
- Smith, A. (1998). Fighting battles, winning wars. *Journal of Conflict Resolution*, 42, 301–320.
- Volden, C., & Wiseman, A. E. (2007). Bargaining in legislatures over particularistic and collective goods. *American Political Science Review*, 101, 79–92.

PART V

POLITICAL THOUGHT

65

THE ANCIENTS

STEVE G. HARTLAUB

Frostburg State University

Western political philosophy is a conversation about the best way to live a good life. The conversation began with the ancient political philosophers Socrates, Plato, and Aristotle. In 1989, with the fall of the Soviet Union, it appeared that the conversation might be over, that the question of the good life had been answered—the good life was the life of liberal democratic capitalism enjoyed by the United States and the other developed nations (Fukuyama, 1992). Most people around the world seem to accept the justice of a democratic government and the goodness of a capitalist lifestyle. Still, one must wonder whether other people at other times also thought that their style of government was just and that the life they were living was the good life.

In the movie *Conan the Barbarian*, Conan is asked, “What is best in life?” He responds, “To crush your enemies, see them driven before you . . . and to hear the lamentation of their women.” How can we be sure that we are right and Conan is wrong? We judge Conan to be wrong because we can look at his society from the outside, from the perspective of our society. To judge most things, it is helpful to get some distance from them—if you zoom out on Google Maps, you can often find your way much easier than you can if you view something at very close range. This is the single most important reason to study ancient political philosophy; it gives us some distance and perspective on our own society. It asks the same kinds of questions, but often has radically different answers.

This distance is, unfortunately, also a small obstacle facing the student of ancient philosophy. The ancients did not write in English; they wrote in ancient Greek or Latin. Most students (and frankly, most professors) do not read ancient Greek and Latin, so we have to rely on translations by other people, but how can we be sure that they are translating the work correctly? If you ever played the game in which you sit in a circle and whisper a sentence from one person to another, you know that meaning is often lost in translation. We will return to this problem below because it relates to a major area of contention within the field.

A second minor difficulty in understanding the ancients is that the words they use are both familiar and strange. When you read Shakespeare in high school, you discovered that even though he wrote in English, it is often not clear what he means. Often words that we use today were used in a different sense, or in a number of different senses, in the past. The same thing is true in Plato (trans. 1991) and Aristotle (trans. 1984): Many of the words they used are familiar to us because they often form the root of words we use today, but they do not mean the same thing. For example, the word *polis*, which is the root of the modern word *politics*, refers to the city-state. Greek society was generally organized around a small metropolitan area, or city-state, that encompassed about 100,000 to 150,000 people and the farmland necessary for their support (Barker, 1960). One of the key

elements of the city-state is that it was self-sufficient—it had everything that a person needed in order to live a good life. When Aristotle argues that man is the political animal, he means that we can become fully human only within a polis.

This is not how we would understand the term *political animal* today. We would take it to mean either someone who loved politics or someone who constantly tries to promote his or her own advantage through persuasion and power. This difficulty is addressed by the frequent use of the Greek word in place of English. When studying ancient political philosophy, students probably need to learn a few Greek words, such as *polis*, that have distinct meanings. Fortunately, the need for new vocabulary is relatively small. The ancients usually start with common opinions about things and discuss them in everyday language. They believe that common opinions contain some truth but that these truths are incomplete and sometimes contradictory. Their normal method is to push the logic of common opinion until we understand its limitations and, we hope, where to go from there.

One last feature of ancient writing is the strange notations in the margins of most editions of the works of Plato and Aristotle (e.g., “454d–e”). These numbers are known as *Stephanus numbers*, after an early collected edition of Plato’s work. They are used today to allow for easy comparison across different translations of a work and to give the reader a more exact reference point. These minor difficulties sometimes discourage students from reading the ancients. This would be a tragic loss. The ancients saw the world differently from the way we modern people do. Exploring their thinking allows us a perspective on our own lives and thought that can scarcely be gained in any other way.

The remainder of this chapter is divided into two main sections. The first section provides an overview of the basic understanding of politics found in the ancient philosophers Plato (trans. 1991) and Aristotle (trans. 1984). Although there are important differences between Plato and Aristotle, this section focuses on the common ground that distinguishes ancient political philosophy from more modern approaches: its understanding of nature, human nature, and the polis. The second section introduces some of the basic academic questions and approaches that relate to studying the ancients.

The Common Ground

Ancient political philosophy is dominated by three related figures, Socrates, Plato, and Aristotle. Socrates was the founder of political philosophy, Plato was his most famous student, and Aristotle was the most famous student of Plato. All three spent major time in Athens during the 5th and 4th centuries BC. During their lives, Athens was an incredibly powerful and vibrant city-state with a flourishing culture. Out of this crucible of intellectual and artistic

creation, political philosophy was created with the recognition of the distinction between nature, convention, and the divine.

Nature

Natural things are the way they are *without* human action or manipulation. *Conventional things* are the way they are *because* of human action or manipulation. Divine things are *beyond* human action or nature; they are supernatural. The main focus of political philosophy is to distinguish what is natural from what is conventional in human beings and society. As Plato (trans. 1991) points out in the *Republic*, it is conventional that women wear their hair longer than men do. In different societies, and even at different times in the same society, the fashionable length of a person’s hair has been long or short, tied back or worn loose, dyed purple or spiked. What is conventional changes over time and place; what is natural is eternal and cannot be changed by society. The fact that women bear children is not a product of our society, and no matter how many movies Hollywood makes about it, men will never be able to give birth. The understanding of nature is one of the major distinguishing features between ancient and modern political philosophy.

There are two related senses in which we use the term *nature*. Imagine that you and a friend are flipping channels trying to decide what to watch on TV. If your friend wants to watch a nature program, you would assume he or she wants to watch a show about walking dinosaurs or marching penguins. This first sense of the word *nature* means the physical external world. While continuing through the channels, you come across a detective show. In the scene, a detective has just burst into the library and found a dead man with a beautiful woman standing over him. The detective asks, “What is the nature of your relationship to the deceased?” The woman could answer, “He was my brother,” “my husband,” “my secret lover,” “my boss,” or many other things. Any one of these descriptions would tell the detective something important about the case. This second sense of the word *nature* tells us the essential character of a relationship or a thing—to know the essential character of a thing is to know its nature. It is the nature of water to flow downhill. Ancient political philosophy focuses on this second sense of nature, the essential unchanging character of human and political things.

To better understand what differentiates the ancient understanding of nature, consider Aristotle’s four causes: the *material*, *efficient*, *formal*, and *final* cause. These can be illustrated by the construction of the three little pigs’ houses. The *material* cause is the physical matter or raw materials that go into each house—the straw, sticks, or bricks. The *efficient* cause is the motion or activity that went into constructing each house—the straw house requires the least work, the brick house the most. The *formal* cause is the plan or blueprint. Each little pig had to have some idea of what the house would look like when it

was done. Last, we have the *final* cause, or purpose (telos), of the house—to keep them safe from the big bad wolf.

The ancients would argue that we can understand the house, including what *caused* the house, only in terms of its plan or purpose. Modern science, conversely, tends to look only at the first two of Aristotle's causes: matter and motion. This is the essence of the Newtonian worldview. Modern political philosophy follows the lead of natural science by focusing on the basic drives, or physical nature, of human beings in order to explain why we act the way we do. This atomistic view is taken up by Hobbes and most of the modern political thinkers who reject any notion of a plan or purpose to human existence; they tend to build their philosophies from the bottom up.

The ancients, on the other hand, thought that the nature of a thing included a purpose or goal that is inherent in the thing itself, its nature. This sounds a bit strange to modern ears, but it does have a certain logic to it. For example, the purpose of an acorn is to grow into an oak. This is the reason acorns exist, their telos. The fact that not all acorns will become oaks does not deny their purpose. The telos of a thing does not guarantee its completion; it provides only the end at which the thing aims. Humans, and hungry squirrels, can prevent acorns from becoming oaks, but they cannot change the acorns' nature. Humans and squirrels cannot make acorns grow into elm trees or pineapples.

The ancients would go further. They would argue that the purpose of the acorn is not simply to grow into any oak but to grow into the best possible oak. In other words, inherent in the idea of an oak is the idea of the best possible, or excellent, oak. The best possible oak is not a utopian idea, where every leaf, every branch, every cell is perfectly constructed. Rather, the best possible oak is what would happen if an acorn were raised in the best soil, with clean water, adequate wind, and no hungry squirrels. This *idea* of the excellent oak is why the ancients are often called idealists. It is important to remember, however, that the ancients' *idea* is not a perfect world but the best that can be attained on this earth with the things that already exist on this earth.

Obviously, not all acorns will become the best oak; in fact, none may ever have achieved this goal, because the right circumstances have never occurred. This does not mean that it is impossible but merely that it is unlikely. The idea of the excellent oak allows us to judge all other oaks, even if we have never seen the best possible oak. We can tell whether an oak is stunted by poor soil, bent by too much wind, or rotting from fungus. The idea of the excellent oak forms a universal and eternal standard by which we can judge every other oak we see (Best, 1997).

Human Nature

Socrates appears to be the first person to take this type of analysis and turn it toward the study of human beings and the city. Philosophy becomes political when it turns to the study of human nature—the essence of what it means to be human. The ancients believed that human beings

have an essential character, or telos, just like an acorn does. For the ancients, human beings are distinguished from other animals by their capacity for reasoned speech, or *logos*, and by the *subject* of their reasoned speech—justice, or what is the good life. As Aristotle (trans. 1984) explains, “It is peculiar to man as compared to other animals that he alone has perception of good and bad and just and unjust . . . and partnership in these things is what makes a household and a city” (1253a1). Lions do not debate whether killing that sickly little zebra will ultimately benefit the zebra population. They act on instinct and kill the easiest target. Human beings, by contrast, organize their social structures by choice, at least when they reach the level of the polis.

Aristotle (trans. 1984) describes the movement from the smallest natural society of men and women to the creation of villages that meet nondaily needs and then to the melding of several villages into a polis, a partnership founded on an idea of justice. The critical aspect of the polis is that it is self-sufficient: It contains everything that a person needs in order to live a good life. Most important, the polis allows us to debate justice and the best way to live. The polis is natural because it is inherent in humans to discourse about justice, and this leads inevitably to the construction of the polis. This is why humans are *the* political animal: because outside the polis, we could not fulfill our natures; we could not be fully human.

A second fundamental aspect of human nature for the ancients is the division of human nature into two parts, the *body* and the *soul*. The body is the physical aspect of our existence, and the soul is the mental aspect. The soul is another word that has many modern connotations that are not entirely meant by the ancients. The Greek word for soul is directly related to the modern word *psyche*, the root of psychology or the study of human thought and behavior. The Latin word for soul is *anime*, the root of the modern word *animation*. So for the ancients, the soul is the source of voluntary movement and behavior, as opposed to the involuntary movements of the body.

The soul is further divided into at least two parts, with a basic distinction made between the *logos*, or reasoning part, and the more emotional part of the soul. The fulfillment of human nature for the ancients is the proper development and balancing of the parts of our nature. The body should be ruled by the soul, not the soul by the body. Our bodies are often hungry or thirsty, but we do not instantly gratify our physical urges by shoving any available food and drink into our mouths as fast as possible. Our soul compels the body toward the proper fulfillment of our urges in the right time and the right measure. Further, the reasoning part of the soul should rule over our emotions or appetites. We are often angry and would like to say something offensive to our boss, or even our professor. Most of the time, our reason can calm our anger, and we can hold back the snide comment or offensive remark. Part of what it means to become an adult is learning to control our emotions and channel them in constructive ways.

Just as the telos of the acorn is to become an excellent oak, so the telos of a human is to become an excellent person, and an excellent, or virtuous, person has a soul that is in harmony, where each part minds its own business. In other words, excellent, virtuous people have a healthy body with normal urges. They have emotions, but at the right time and the right place and for the right reasons. Finally, they are reasonable, debate justice and the right way of life, and have the ability to live by what they decide to be correct. Aside from this basic ordering of the soul, there are specific virtues, such as courage, moderation, and justice, that are a major part of the debate about human excellence in ancient thought. One thing modern students often find frustrating about the ancients is that they do not provide simple answers to the moral questions of life. The moral questions of life are unique and require that we consider them within our political community. Ultimately, good moral choices require us to be raised in a way that allows us to be good moral people, and this can occur only within a polis.

The Polis

The central political unit for the ancients was the polis, and only through understanding the polis can we understand human nature. Our nature is never directly accessible to us because we are all raised within a particular society—the conventional laws under which we were raised and live always condition our underlying human nature. To understand the polis, the ancients looked at the constitution of the polis—the fundamental values and principles that form a particular society. The U.S. Constitution, for example, identifies the fundamental values and principles of the country, those of the Declaration of Independence, and then goes on to construct institutions and distribute power in the way thought best to secure those principles. Most countries do not have written statements of their principles, but their principles can be discerned by looking at who rules the society.

In both the *Republic* of Plato and the *Politics* of Aristotle, the question of who should rule is a central focus. Those who rule imprint their values on the society through their example and through the laws. For the ancients, laws are more than a statute or legal pronouncement; they are the fundamental values that underlie and bind a society together. The rulers embody the principles of their society by possessing those characteristics that are revered or valued by society. In an oligarchy, for example, it is the wealthy who rule because the society is primarily concerned with the accumulation of wealth. The creation of wealth demonstrates the characteristics that the society values—not simply the wealth itself, but hard work, savings, management ability, and the other things that allow people to gain wealth.

A number of groups contend to rule, with the three most prominent being the *many*, who support democracy; the *few wealthy*, who favor oligarchy; and the *few virtuous*, who favor an aristocracy. According to Aristotle (trans. 1984),

these groups each “fasten on a certain sort of justice, but proceed only to a certain point, and do not speak from the whole of justice in an authoritative sense” (1280a1). Some, like the oligarchs, believe that because they are unequal in wealth, they should be unequal in rule, that is, they should have more of everything. The many, or demos, believe that they are equally free citizens, and, therefore, they should have an equal share in ruling. The central question is “equality in what sort of things and inequality in what sort of things . . . for this involves a question—and political philosophy” (Aristotle, trans. 1984, 1282b1: 21–24). Here we have the essence of political philosophy, deciding what kinds of human equality and inequality matter and then figuring out how they should be incorporated into the society.

Ultimately, the ancients resolve the question of who rules in favor of those who possess the most virtue. Those who have the most knowledge and the best moral character are most likely to make the best decisions for the community as a whole. The ancients generally believed the rulers should be granted whatever power is necessary to make the best city. In order to produce the best possible oak tree, we would need to control its environment, the soil, water, sun, and wind. If the goal of the polis is to produce the best people, the ancients argue that we have to control the social environment of human beings—family, wealth, education, even entertainment and music. Only in the polis can the environment be constructed in such a way that people will be directed toward fulfilling their nature: the complete practice of virtue. If people receive a proper education, they will internalize the values of the regime, and there will be little need for written laws and punishments. Much of Plato’s *Republic* (trans. 1991), for example, is an effort to construct a hypothetical city in speech in order to find out the truth about justice and the just life. In the course of this discussion, the characters in the *Republic* determine what stories people should hear, what music they should listen to, how they should exercise, whom they should marry, and numerous other things we would view as private.

This understanding of nature, human nature, and the polis provides the common ground of ancient political philosophy (Best, 1997). This broad agreement hides the rich diversity of opinions and perspectives that exists between Plato and Aristotle and among other great writers of the ancient world: Xenophon, Cicero, Thucydides, and Saint Augustine. The depth and diversity of ancient political philosophy are why succeeding generations have returned to it again and again and why the study of ancient political philosophy is still a major part of political science.

The Study of Ancient Political Philosophy

The academic study of ancient political philosophy began almost as soon as ancient political philosophy. From the very beginning—because Aristotle disagreed with

Plato—there has been a sustained debate over the question of the good life that has provoked new participants to join the conversation. The conversation was alive through the rise of Rome in the writings of Cicero and at the birth of Christianity in the writings of St. Augustine. It went almost silent in the West during the Dark Ages, only to be revised by the Renaissance and the writings of modern political theorists. With the birth of modern political science in the past century, the conversation over the good life has continued. The rise of science has generated three problems that confront the modern student of ancient political thought: the problem of *authenticity*, the problem of *knowledge*, and the problem of *truth*. Despite these problems, or perhaps because of them, a number of current streams of research can be identified.

The Problem of Authenticity

One aspect of ancient political philosophy that is different from most modern political philosophy is the debate over the authenticity of the texts. There is virtually no debate over which books or pamphlets were written by modern political theorists such as Machiavelli, Hobbes, Locke, or Rousseau. Conversely, the ancient political theorists, who wrote thousands of years ago, have provoked a wide and extensive debate on the authenticity of the texts. Approximately 36 dialogues and a few letters are attributed to Plato. We have all Platonic dialogues discussed in ancient writings, but we cannot be entirely certain that Plato in fact wrote all of them. Starting in the 19th century, there has been an extensive academic debate over the authenticity of the dialogues (Irwin, 2008; Taylor, 1964). What we have of Aristotle's writings are much less complete. Aristotle is known to have produced numerous dialogues and other finished pieces, but all these have been lost to history. What we have of Aristotle's writings seem to be collections of lecture notes. There is even some debate whether they are Aristotle's notes or a student's or colleague's (Barker, 1948; Lord, 1984).

A second major debate within academia is the attempt to trace the development of a philosopher's thinking through the timeline of his writing. It is felt that Plato, for example, changed his thinking over time and that we can understand the development of his ideas over the course of his life by arranging the dialogues in the order he wrote them. For example, the *Laws* is widely considered to be Plato's last work, and yet some of the ideas presented in it seem to disagree with elements of the *Republic* or other works. This difference might reflect the development of Plato's thought, or it might be a different presentation of the same basic ideas to a different audience (Taylor, 1964).

The Problem of Interpretation

There are a number of ways to read the texts of political philosophy. At first the issue seems straightforward: Just read what was written, and then you will understand. But

this is not always the case. Let us imagine two people, John and Jane, who have been dating for a couple of months. One day, John senses that something is amiss, so he asks Jane what's wrong. She responds, "Nothing." Should John take this at face value and go out with his friends, or should he consider the context of her remarks? Are Plato and Aristotle telling us the truth in their writings, or do we have to consider the context in which they wrote to understand their thinking?

A second problem immediately emerges: Which of the many opinions put forth by Plato, and to an extent Aristotle, represents what they really thought? Plato wrote almost entirely in dialogues, which read more or less like plays, but Plato himself is never a character. In most of the dialogues, Socrates is a central character, and we are left with the task of separating the character of Socrates from the thought of Plato. Aristotle, on the other hand, frequently explores numerous different positions and rarely comes down with hard and fast rules for behavior. If their intent was to convey their thinking to the future, then dialogues (or lecture notes) seem a confusing way to do so.

Most readers believe that the format of the dialogue was chosen specifically to avoid simply telling people what to think. The Socratic and Platonic idea of education is that it takes place through *dialectic*—reasoned conversation. It would be counter to this idea to write down a set of truths for people to memorize and recite. They would not *learn* them but only memorize them. In the *Wizard of Oz*, Dorothy always has the power to return to Kansas, but she can do so only after she has learned some things for herself. In the same way, Plato is trying to teach us something in the dialogues, and this means he cannot just give us all the answers; he has to develop our reasoning and point us in the right direction. In a similar way, Aristotle's two main political works, the *Politics* and the *Ethics*, do not reach many hard and fast conclusions. Rather they mull topics over, considering them from different perspectives at different times. These ancient philosophers do not want us to simply follow their teachings; they want us to learn the truth for ourselves.

The Problem of Truth

The central debate over how we should interpret the ancients is conditioned by a larger debate within the social sciences over the nature of truth. The tradition of political philosophy is a search for unchanging truth, but we are faced with an immediate problem: Philosophers each claim that they have discovered some eternal wisdom about human nature, society, and the good life, and yet their answers do not agree with one another. These divergent views on the truth of human life have resulted in two major reactions regarding the idea of truth. On one side are those who still believe that there is eternal wisdom about human nature and the polis and that these truths can be discovered by human reason and used to guide human life. They argue that we need to join the conversation and work

to separate truth from error and discover for ourselves the essential truth of a good life—in other words, to become philosophers.

The second major reaction to the diversity of viewpoints is to recognize that each society has produced a set of values that define truth for that society. Human beings naturally produce a set of values that define right and wrong, and these can be understood only on their own terms in the context of the society that produced them. These “truths” make sense within a particular social structure, but they do not represent truths that transcend time and place; they are not eternal truths. The idea that all people and ideas are a product of the social, political, and economic systems of which they are a part is called *historicism*.

Of course, to be honest, there is a third reaction: give up. Most people choose to stop reading political philosophy, accept the truth that society gives them, relax in their La-Z-Boy, turn on the TV, drink beer, and eat pork rinds. The vast majority of people are satisfied with the answers they are given and choose not to question them. There are always a few, however, who want to know more, who do not accept the truth of their situation without question. The careful study of political philosophy provides these few a way of waking themselves from the all-too-comfortable daze of modern society.

The question of how we should interpret ancient writings is conditioned by whether one believes in the possibility of eternal truth or in historicism. Those who believe in the possibility of eternal truth generally focus very closely on what the authors have literally written in the texts themselves. The assumption is that Plato and Aristotle wrote their works to communicate their ideas to others, so to understand their ideas, we need to focus on exactly what was written. This perspective is very concerned with the accuracy and consistency of translations to ensure that nothing comes between the thought of the ancients and our minds. The majority of scholars who accept the possibility of truth read Plato and Aristotle in this way, as if they always say what they mean (Barker, 1948; Taylor, 1964).

An offshoot of this approach to understanding the text, developed primarily by Leo Strauss (1988), is to consider whether Plato and Aristotle might have hidden some of their most provocative ideas between the lines. Strauss believed that the search for truth always requires calling into question some of the accepted truths of the current society. Philosophy, as the search for truth, is therefore always under the threat of persecution from the authority of the state because it must challenge the fundamental beliefs that underlie that authority. This persecution “gives rise to a peculiar technique of writing, and therewith a peculiar type of literature, in which the truth about all crucial things is presented exclusively between the lines” (p. 25). This kind of literature, of which Plato and Aristotle are examples, is divided into an *exoteric*, or surface, teaching and the *esoteric*, or hidden, teaching. The problem of interpretation is then trying to read between the lines to

discover this hidden teaching. The answers are still within the text, but they are hidden because, according to Strauss, “There are basic truths that would not be pronounced in public by a decent man, because they would do harm to many people who, having been hurt, would naturally be inclined to hurt” the philosopher in return (p. 36).

Interpreters who believe that truth can be understood only in the context of the society are less concerned with what the text says and more concerned with the social and political context in which it was written. Scholars from this perspective are less concerned about the accuracy of translations, because there is no real need to analyze and decipher the literal, or hidden, meanings of the text. Rather than pursuing some eternal teaching on truth in the writings of ancient philosophers, the reader needs to interpret their writings in light of the historical context in which they wrote. For example, to understand Aristotle’s justification of virtue, we must understand how this understanding of virtue fit into the economic and social structure of Athens at the time. Only through an analysis of the social and political system can we understand how this truth fits into the system of thought that was ancient Athens.

Recently, an offshoot of this approach has developed that can be called *contextualizing* (Richter, 2009). Whereas historicism interprets the ancient writers by looking at the social and political context in which they wrote, contextual interpretation tends to focus on how the ancient philosophers have been *read*. Different epochs in human history have read Plato and Aristotle differently. Sometimes they are read as essentially democratic, other times they have been read as supporting kingship and monarchy. A contextual approach focuses on the context of the readers more than the context of the writings: Contextualists look at how different contexts influenced the interpretations of the texts. For example, some feminists would argue that the tradition of Western political philosophy is biased, not only because it focuses almost exclusively on the writings of men but also because it has been interpreted only by men (Saxonhouse, 1985; Shanley & Pateman, 1991).

The Next Step

There is no substitute for reading the writings of Plato and Aristotle, available in many editions, for yourself. Philosophic commentaries are like movie reviews. They can help us notice some things we might miss, but they are not a substitute for seeing the movie ourselves. Two of the most accessible Platonic dialogues are the *Apology* and the *Crito*. Each is fairly short and considers a limited range of issues. The *Gorgias*, although not as popular in the discipline, is a good intermediate step in preparation for the *Republic*. The *Gorgias* follows a fairly straightforward narrative structure and contains a series of interconnected arguments on rhetoric and the best life. The *Republic*, on the other hand, considers a wide array of topics within a fairly complicated narrative structure. Although the *Republic* is a tremendously rewarding work, it is also quite

intimidating and requires careful consideration and effort to work through. Aristotle's two principal political works are the *Politics* and the *Ethics*. These works complement one another nicely although the *Politics* is more directly of interest to political scientists.

Students who tend to like a more creative style of writing often find Plato more accessible, while students who prefer a more logical, systematic exposition often gravitate toward Aristotle. The study of any of these works is best done in a class or in a reading group, as it is very useful to have other people to test your ideas and interpretations on. If this cannot be done, you can read a good interpretive essay along with the work. The long interpretive essay that accompanies Allen Bloom's translation of the *Republic* (Plato, trans. 1991), for example, provides a lot of food for thought (see also White, 1979). Mary P. Nichols (1992) provides a solid exposition of the *Politics*. Whether you agree or disagree with it, a good interpretation compels you to read and understand the work more thoroughly. For a general overview of the subject, *The Mainstream of Western Political Thought* is a very clear and readable overview (Best, 1997). For a more thorough treatment, the *History of Political Philosophy* (Strauss & Cropsey, 1987) contains short interpretive essays on almost every philosopher of note.

Conclusion

This chapter has outlined the basic foundation of ancient political philosophy and how modern scholars approach its study. The different understandings of nature, human nature, and the polis developed by the ancients provide an excellent antidote to the parochialism of our time. The ancients raise questions that challenge many of our fundamental beliefs and thereby make us defend them, turning our blind prejudices into reasoned arguments and beginning the process of education.

Within the discipline of political science, the need to challenge our fundamental beliefs is no less critical. Our basic understanding of the major texts of Plato and Aristotle is constantly being challenged by new interpretive methods and approaches. In addition, there has been increasing interest in exploring less-well-known thinkers of antiquity such as Xenophon, Augustine, and Cicero. Another approach is to focus on a particular theme as opposed to a particular work. The thinking on justice, slavery, economics, democracy, and many other topics is increasingly considered across texts or theorists. Yet another approach is to try to apply the thought of the ancients directly to the modern world to shed light on modern problems (Koivukoski & Tabachnik, 2005).

The study of ancient political philosophy is partly directed at understanding the development of Western political thought through time. Whether one generally agrees or fundamentally disagrees with the ancients, their thought is critical to beginning our ascent from the parochial

views that surround us. More important, the ancients considered the eternal question of how one leads a good life. Should I be moral or immoral? Should I pursue a life of physical pleasures, or are the pleasures of the soul more valuable? These questions are front and center as students work through their education, even though many classes do not seem to try to answer them. Ancient political philosophy speaks directly to the problems of living a good life, and students who study it are rewarded with a clearer understanding of these problems and, perhaps, their solutions.

References and Further Readings

- Aristotle. (1984). *The politics* (C. Lord, Trans.). Chicago: University of Chicago Press.
- Barker, E. (1948). *The politics of Aristotle*. London: Oxford University Press.
- Barker, E. (1960). *Greek political thought*. London: Methuen.
- Best, J. A. (1997). *The mainstream of Western political thought*. Lanham, MD: University Press of America.
- Fukuyama, F. (1992). *The end of history and the last man*. New York: Free Press.
- Hartlaub, S., & Jelier, R. (2004). *Building a community: Ancient political philosophy and new urbanism*. Paper presented at the meeting of the Midwest Political Science Association, Chicago.
- Hartlaub, S., & Lancaster, F. (2009). *Taming the tyrant: Aristotle and Machiavelli on tyranny*. Paper presented at the meeting of the Midwest Political Science Association, Chicago.
- Irwin, T. H. (2008). The Platonic corpus. In G. Fine (Ed.), *The Oxford handbook of Plato*. New York: Oxford University Press.
- Koivukoski, T., & Tabachnik, D. E. (2005). *Confronting tyranny: Ancient lessons for global politics*. Lanham, MD: Rowman & Littlefield.
- Lord, C. (1984). Introduction. In Aristotle, *The politics* (C. Lord, Trans.). Chicago: University of Chicago Press.
- Nichols, M. P. (1992). *Citizens and statesmen: A study of Aristotle's Politics*. Savage, MD: Rowman & Littlefield.
- Plato. (1991). *The republic* (2nd ed., A. Bloom, Trans.). New York: Basic Books.
- Richter, W. L. (2009). *Approaches to political thought*. Lanham, MD: Rowman & Littlefield.
- Saxonhouse, A. W. (1985). *Women in the history of political thought*. Westport, CT: Praeger.
- Shanley, M. L., & Pateman, C. (1991). *Feminist interpretations and political theory*. University Park: University of Pennsylvania Press.
- Strauss, L. (1988). *Persecution and the art of writing*. Chicago: University of Chicago Press.
- Strauss, L., & Cropsey, J. (Eds.). (1987). *History of political philosophy* (3rd ed.). Chicago: University of Chicago Press.
- Taylor, A. E. (1964). *Plato: The man and his works*. Cleveland, OH: World.
- West, T. G., & West, G. S. (1984). *Four texts on Socrates*. Cornell, NY: Cornell University Press.
- White, N. P. (1979). *A companion to Plato's Republic*. Indianapolis, IN: Hackett.

66

ASIAN POLITICAL THOUGHT

TIMOTHY J. LOMPERIS

Saint Louis University

Scholars of Western political thought have not disputed the fact that there is a rich body of political thought in Asia. They have just not bothered to incorporate it into their corpus. This chapter seeks to provide long-overdue recognition to this body of thought by calling attention to the fact that despite its heavy religious content (until modern times), the encounter with political ideas in Asia is just as profound as it is in the West. In fact, since these ideas in Asia are heavily fertilized by their Western colonial legacy, the West has much to learn about itself from these Asian borders to the West's material and intellectual reach.

In this presentation of Asian political thought, what will emerge is that the such central ideas as *democracy*, *freedom*, and *equality* were formed in a historical context different from the West. In the West, these ideas were expressed and then refined through a prism of small city-states in Greece, the universal empire of Rome, the subsequent collapse of this imperium politically but its persistence intellectually in the Thomist medieval synthesis, the smashing ferment (both intellectually and institutionally) of the Renaissance and the Reformation, and the birth of the modern nation-state in the twin crucibles of the Thirty Years' War (1618–1648) and the French Revolution (1789–1795).

In Asia, these same ideas have been definitionally filtered through a different historical stage in a play of three acts. The first act is the traditional or classical era before the Western contact. We will see what from this period endures as a mark today of "Asianness." The second act is

a scrutiny of the trauma of the colonial experience. The vast majority of Asian societies, either directly or indirectly, came under Western colonial control or under spheres of Western influence. How to react to this intrusion precipitated a major crisis but also resulted in a rich intellectual ferment that produced the first articulations of Asia's nationalisms. The third act is the modern period from the end of World War II to the present, when Asia was set free on its own independent course. This has raised the question, Whither modern Asia? Is Asia no different from a common globalizing world, or does something distinctively Asian remain about its political thought?

In these three acts, we will examine Asian concepts of the state and of statecraft, as well as of military grand strategies and views on social equity and gender as they relate to these three concepts. The focus will be on India and China because these two ancient polities form the foundational pillars of Asia. Japan will also be given considerable attention, along with some references to Korea. Southeast Asia will be considered not so much as individual countries but as a region that has always been a tempestuous battleground between Indian and Chinese ideas and institutions.

Theoretical Approach

Insofar as the political thought of Asia came to the attention of Western political theorists, it tended to be painted in the broad brushes of overgeneralization. Karl Marx, in

outlining the global stages to his class struggle, wrote of an “Asiatic mode of production” (quoted in Tucker, 1972, p. 5), which he characterized as a labor-intensive agricultural society. Writing in this tradition, Karl Wittfogel (1957) spelled this out as a form of “Oriental despotism” arising from the need to secure the necessary *corvée* labor to support the rice culture of what he termed “hydraulic society.” Taking a more cultural perspective, F. S. C. Northrop (1946) distinguished Asia as having a more aesthetic *weltanschauung* than the scientific West. Rather than the clear subject-object divide in the West, Asia, Northrop contended, charted reality along a more fused aesthetic continuum, thereby creating different logics and perceptions about the world.

More recently, such political scientists as Lucian Pye (1985) and Daniel Bell (2000) have remarked on the different conceptions Asians bring to politics. To both, these differences require democracy, in particular, to undergo considerable modification for any successful transplantation to Asia. For Pye, the changes will have to allow for a more dependent and paternal understanding (and acceptance) of power. And for Bell, for Asia to be comfortable with democracy, democracy will have to give a special place to knowledge over and above mere democratic egalitarianism.

This is because ideas of democracy, freedom, and equality have developed out of a historical context different from the West's. This context has led to conclusions on the grounding of these ideas that are also different from the conclusions of the West. Put simply, Western political thought is grounded in the individual as the basic unit of politics, and in equality, in some form, as the accepted basis for human relations and political rule. In the Asian context, political thought came to be grounded in the group, not the individual, and in hierarchy, not equality. As shall be clear from the description of the context of three historical acts, the contact of the ideas of democracy, freedom, and equality with Asia calls for some reformulation. In fine, this chapter explains that in passing these three ideas through an Asian historical encounter, one can arrive at richer, multicultural definitions of such seemingly universal political ideas.

Classical Asia

Asia has provided an arena for all the world's value systems. Hinduism is the oldest. Its earliest forms were similar to the religion and ideas of the ancient Greeks. Perhaps the Indo-Aryan invaders of the Indian subcontinent effaced the same Triple Goddess overrun by Jason and his Greek Argonauts in the Black Sea city of Colchis. In any case, Hinduism emerged in the first millennium BCE as a religion and political culture of conquest. Buddhism arose later as a sort of Lutheran reformation to Hinduism. It held distinctly gentler political ideas. This gentler faith, however, was literally obliterated by Muslim

invasions into the subcontinent that began in the 8th century CE. (Buddhism went on to thrive in China, Japan, Korea, and Southeast Asia.) These new invaders oscillated between two approaches in their new dominions. One was to exterminate opposition and force Islam by the sword. The other was to cooperate with local power groups and rule by accommodation. As it spread to Southeast Asia, Islam became more moderate and diffuse in its ideas and practices.

In China around the 6th century BCE, Confucianism developed its own order among society, nature, and the cosmos. This ordering principle, of the dual forces of yin and yang, was an early portrait of a historical dialectic similar to that in the writings of Heraclitus, Hegel, and Marx. While Confucianism propounded a rigidly hierarchical sociopolitical order, the “turning of the wheel” from Buddhism and the “reversion of the Dao” from Daoism introduced the idea of reciprocity. Mencius politicized the role of the emperor by entrusting him with the Mandate of Heaven, but in tying this mandate to reciprocity, Mencius also gave the people the right of revolution. Daoism added the mystical and the magical to this mix. For all its order, this ancient Chinese system gave birth to a romance of protest, with sage-knights acting as Robin Hoods. These folk heroes later inspired modern revolutionaries such as Mao Zedong (Schwartz, 1985).

In this Asian drama, as in Europe, there has been a gradual growth of secularism. But modern secularism has never been completely successful in India, and religion has never died in China. In India, religion represents a complete value system. This heavily religious value system, however, did not preclude lengthy and systematic treatment of political questions. The epic *Mahabharata* contains long political essays on statecraft, kingship, and military strategy. One ancient text, Kautilya's *Arthashastra*, introduces all Machiavelli's ideas about political survival more than a thousand years earlier than *The Prince* (Basham, 1959). China demonstrated a more robust tradition of secularism, partly because Confucianism never really addressed the question of God. Buddhism filled this gap. The Legalists attempted to place law as a higher principle of social ordering than cosmic rhythms of yin and yang. But dynastic rulers preferred the ambiguities of the cosmos to the concrete constraints of the law. In China, too, as in all Asia, religion stayed on top, fusing society and politics to the sanctity, sanctions, and political protection of the gods (Schwartz, 1985).

More than on top, the Yamato clan in Japan proclaimed themselves to be gods. In their success, they have provided Japan with the longest single line of kings in world history and a sense of nationalism and ethnic identity that runs very deep. Although “divinely” ruled, the Japanese never saw themselves as holding the gateway to heaven. They were, then, not averse to borrowing, and they looked to Confucianism and Buddhism to order their state and meaning system. Ironically, integrating this borrowing into indigenous Shinto beliefs became men's work. The further

development of Japanese culture—its novels, ceremonies, and haiku poetry—was left to the creative talents of women. Although gods reigned, warriors ruled and warred in Japan. A strong knightly code of Bushido steered the ruling samurai class in the political culture of the warrior-ruler-knights (Yuzan, 1941).

Meanwhile, great kingdoms arose in Southeast Asia, mostly on borrowed Hindu ideas transmitted by Theravada Buddhism from Sri Lanka (Ceylon). There was the Kingdom of Ten Thousand Elephants in Laos, Borobuddur and Bali in Indonesia, and the Khmer empire in Cambodia. The latter's capitol, Angkor Wat, is still the largest religious building complex ever built. Political ideas and institutions in this porous, vulnerable region were mostly Indian (the Chinese influences in Vietnam were the notable exception), but the societies of much of Southeast Asia were ethnically Malay and were held together mainly by their customary *adat*, or customs. These customs set up three social classes (a ruling aristocracy, free land holders, and slaves) bound together in a network of mutual obligations and responsibilities. In this *adat*, property and authority could be held and inherited just as easily by women as by men. When the Muslims came to Southeast Asia in the 13th and 14th centuries, they had about run out their political tether and lacked the vehemence that they displayed in India. They superimposed the veneer of their sultanates on Malaya and Indonesia but were content to have the sultanates upheld by Hindu and Buddhist political principles and by the Malay social *adat* (Tambiah, 1976).

In classical Asia, then, politics were decidedly authoritarian, and more specifically regal, rather than democratic. In India, nevertheless, besides just guaranteeing order, or *danda*, kings were obliged to promote the welfare of the people. In China, this promotion extended to the principle of reciprocity and even to the right of the people to rebel. Nevertheless, freedom in classical Asia was more of a religious goal than a political right: freedom from the cycle of rebirths in India and in the cultivation of an inner peace of the soul in China. Thus, in both societies, freedom was a private preserve separate from the crush of public (communal, religious, and political) responsibilities and duties. In these feudal systems of Asia, these responsibilities were mainly to hierarchically ordered groups. Equality, then, was a relative value and was tied to the status and position of one's group compared with others. Any equivalence to modern Western ideas of equality could be procured only within one's group (and primarily for one's family), not outside it.

Colonial Asia

The conquests of Western imperialism shattered this order. Most of Asia was directly colonized. Even those who escaped direct rule—like the Japanese, Koreans, Chinese, and Thai—were still pulled into an international political

and economic system dominated by Western imperial powers. Because Asian polities had unbroken institutional histories for two millennia (in some cases), punctuated by their own moments of glory, the question of how to both accommodate and account for this Western imposition and superiority provoked deep soul-searching among Asians.

Nowhere was this more deeply felt than in India, which became the crown jewel of the British Empire of 50 colonies worldwide. Some Indians embraced Western civilization. The British Viceroy, Lord Thomas Macaulay, was partially successful in creating “a class of persons, Indian in blood and colour, but English in taste, in opinion, in morals, and in intellect” (Spear, 1961, p. 257). Later, these scions were called “Brown Sahibs.” In furtherance of this strategy, the British invested in a modern university system for India. A proud accomplishment of this system was the Nobel Prize for Literature in 1913 won by the Bengali intellectual Rabindranath Tagore, writing in the King's English (Metcalf, 2001).

Following in the wake of the British raj were legions of Christian missionaries who preached their “good news” and practiced their social gospel with institutions of social reform. Beyond a nationwide network of schools, they set up hospitals, orphanages, homes for widows, leprosariums, demonstration farms for peasant laborers, and social services for outcasts. Many Hindus, although leery of the “good news,” eagerly took up this cause of social reform and, in the Brahmo Samaj of the 19th century, launched their own social gospel of reform of some of the ills and neglects of Hinduism. Muslims displayed a split reaction to the Empire. Since they were India's previous rulers, some resisted, and they went down to defeat in the Mutiny of 1857. Others, such as Sir Sayeed Ahmad Khan, articulated a path of accommodation with the British, insisting that Islam had no objections to at least the political culture of the West. Indeed, as a monotheistic “religion of the Book,” Islam was the more natural ally of this culture than was polytheistic Hinduism. Still others were not so sure of either the Hindus or the British (Pye, 1985). It was Mohammed Iqbal—poet, theologian, and political theorist—who gave eloquent voice to a separate destiny for Muslims in the subcontinent (Malik, 1971).

Although never a directly ruled colony, the reaction in China was equally intense. Tiananmen Square in Beijing was an architectural declaration that it was the gateway to Heaven. British gunboats brought a string of military humiliations that shattered this gateway. A man who dreamed that he was the younger brother of Jesus Christ proclaimed a new portal and led the bizarre Taiping Rebellion of the 1850s and 1860s. The movement also preached equality for women and, at first, democracy. In its suppression, it might have been dismissed as one of those oddities of history, were it not for the subsequent influence the rebellion had on Mao Zedong and other revolutionary modernizers (Ogden, 2002).

Meanwhile, the Qing Dynasty, China's last, made earnest attempts at reform. Western education replaced classical texts for imperial civil service examinations. Principles of constitutional democracy and parliamentary elections were introduced, as were modern railroads, military academies, and financial institutions. In 1911, the mixture of protest and reform exploded into a nationalist revolution and a nearly 40-year interregnum of chaos. Intellectually, the boiling cauldron of this ferment was known as the May Fourth Movement. In the humiliation of the demands of the upstart Japanese for the Shandong Peninsula at the Peace Conference at Versailles in May 1919, Chinese intellectuals desperately cast about for a prescription for modern power: in the pragmatism and liberalism of John Dewey and the United States, in the militarism from Germany and Japan, in language reform and mass education, in physical culture and the emancipation of women, in the assassinations and communes of anarchism, and even in the communism of Karl Marx and the Bolshevism of Russia (Zhou, 1960).

There was ferment in Southeast Asia as well. Peasants, in a series of protests after World War I, decried the collapse of a traditional social and political order guaranteed by a royalty and feudal retainers that used to safeguard their livelihoods and provide a sense of place and security by the Mandate of Heaven (in Vietnam), the will of Allah (in Malaya and Indonesia), the mandala pattern of politics and international relations (in Thailand and Cambodia), and a transferal of merit from Buddha (in Burma and Laos). After an initial, if reluctant, accommodation with Western power and political institutions, these peasants and emerging intellectuals searched for their own terms of modern survival. The Cao Dai sect in Vietnam, which worshipped an all-seeing cosmic eye as interpreted by Victor Hugo, Jesus Christ, Confucius, Lao Tzu, and Joan of Arc, illustrated this perplexity. The mood of resignation to these confusing, but powerful, outside forces was captured by the popular 19th-century epic poem in Vietnam, *Kim van Kieu*. This poem was a creative remake of an old Chinese story of a filial daughter who stays true to her undeserving father in a life of untold suffering but steadfast devotion. These peasant protests, then, grew out of frustrations over their devotion to a traditional structure that could no longer secure this order (Kershaw, 2001).

In Japan in 1853, the commercial visit of the U.S. naval commander Commodore Matthew Perry found the Japanese at a moment in their history when they were ready for an opening from the outside. Their mature feudal order had reached a point of stagnation. A knightly class of samurai undergirded an aristocracy that held the emperor hostage, even as this monarchy as an institution provided continuity, identity, and a sense of cosmic place for all Japanese. In the name of restoring the emperor to real power (*sonno joi*), aristocratic modernizers overthrew this samurai-dominated regime in what was called the Meiji Restoration. The Meiji Constitution established a

liberal parliamentary system in the name of the emperor. But for all this constitutionalism, the Japanese actually modernized through a military path of war with China first (1895) and then Russia (1905; Gluck, 1985). Along with these impressive manifestations of modern power, the continued hold of samurai values, for all this Meiji "liberalism," was nurtured by the education of all Japanese school children in *The Story of the 47 Ronin*, in which final loyalty was still given to extreme professions of honor, in the name of the emperor. It was a path that tumbled Japan into World War II, its greatest national disaster (Benedict, 1946).

The ferment touched off by European imperialism in Asia was not exclusively one way. Europeans who had prolonged contact with Asian societies were often surprised at what they saw. Despite their political weaknesses, these societies revealed sophisticated and well-articulated cultures. A host of scholars called "Orientalists," many of whom had served as colonial administrators, began to translate back for European audiences the "pearls of the Orient": the philosophic *Upanishads* and the twin epics, the *Mahabharata* and the *Ramayana*, from India, and the *Analects of Confucius* and the *Dao de Jing* of Lao Tzu (Lao Tzu) from China. The most ambitious of these projects was the 19th-century "Golden Bough" series of translations into English, sponsored by Harvard University, of most of Asia's finest traditional works. This impact, however, was more than just informative. Ideas from these translations worked their way into the transcendentalism of the New England *literati* (particularly on Ralph Waldo Emerson's "oversoul"), as well as into the philosophic systems of Martin Heidegger and Friedrich Nietzsche and even into the novels of Herman Hesse, among others (Clarke, 1997).

Unfortunately, some of this romantic "orientalism" turned perverse. In this discovery of the deep cultural roots of Asia, some Western scholars, particularly German, began to see themselves as descendants of an elite Indo-Aryan brotherhood that extended from the Indus River to the Rhine (Muller, 1919). German national socialism subsequently appropriated the ancient Hindu symbol for universal brotherhood as the centerpiece to its flag, the swastika.

At first flattered by this attention, modern Asian intellectuals for their part began to resist this characterization of a separate orientalism as tantamount to a civilizational dismissal similar to the "separate but equal" legal doctrine in the United States that served to perpetuate racial discrimination. Whether intellectual traditions produce culturally distinct ideas or whether universal ideas form and recombine themselves around different intellectual traditions is a pervasive issue of epistemology. For the study of political thought in Asia, however, the unfortunate effect of orientalism has been to dismiss political thought in Asia as being too closely tied to religious constructions to be worthy of secular analytical scrutiny.

Modern Asia

World War II (1939–1945) brought disaster to Europe. Even in victory, the power of Britain and France collapsed, and, with that collapse, their empires unraveled and their hold over Asia ended. In independence, not always easily gained, Asia was now free to find itself and define politics in ways authentic to a free Asia and to the particular set of traditional legacies and aspirations of each of its societies. In this mix of the traditional and the colonial, what set of political ideas and institutions would serve independent Asian nations still having to fend for themselves in an international system of Western creation and continued dominance? In Asia's postwar trajectory of growing economic prosperity and rising global political influence, answers to this question have produced rich and innovative contributions to the ongoing development of political thought per se.

After World War II, all of Asia wanted to regain what Asian countries saw as their lost importance in the world. Jawaharlal Nehru, the first prime minister of India, expressed these hopes for all Asians when, in his exultant Independence Day speech on August 15, 1947, he declared, "Long years ago, we made a tryst with destiny, and now the time comes when we shall redeem our pledge" (cited in Hardgrave & Kochanek, 2000, p. 53). Colonialism, he argued, had drained the wealth and energies of Asia, and now it would just flow back (Nehru, 1959). Although it certainly did not flow back right away, in the opening years of the 21st century, this tryst with a recaptured Asian global importance seems well within reach.

The Indian subcontinent, however, has been plagued by serious differences both as to how to attain an independent India and as to what it would look like. The towering figures in this agony were Nehru and Mahatma Gandhi. Gandhi was the moral father of modern India. After traveling around India for 4 years after his return from South Africa at the age of 41 in 1915, Gandhi discovered his three themes of *poverty*, *unity*, and *independence*. As he made the continuation of British rule untenable, he worried about an India "in pursuit of Lakshmi" (wealth), freed from the moderating restraints of religion. Thus, even as he determined to entrust the future course of India to Nehru, he was troubled by the younger man's Hamlet-like agnosticism (Gandhi, 1957).

Nehru epitomized Macaulay's "Brown Sahib," and Nehru's highly cerebral autobiography, *The Discovery of India* (1946/1959), was really an articulation of his own divided soul. His professed admiration for the ancient Hindu scriptures and epics was profoundly philosophical and somewhat idealized. He preferred to highlight the moments of unity and power and gloss over the divisions and wars of India's past. He could not bring himself to take this philosophical appreciation to a spiritual awakening. For Nehru, the influences of a secular English liberalism

were too strong for this. To him, the best of India lay in its moments of unity around a *chakravartin*, or universal emperor, such as Ashoka, Harsha, or Akbar. Because of India's deep religious and social divides, Nehru felt that this unity could come, in modern times, only under a secular India united by Western principles of liberal democracy. The Congress Party was founded with this as its core credo. Unfortunately, Nehru dulled his economics by embracing the socialism of the British Fabians and the Russian Bolsheviks (he expressed a continual admiration for the accomplishments of the 5-year plans of the Soviet Union). Under Nehru's leadership as prime minister (1947–1964), Lakshmi, the goddess of wealth, remained aloof (Nehru, 1946/1959).

Although Gandhi and Nehru were the giants, other voices arose in the subcontinent. Ironically enough, Mohammed Ali Jinnah, the father of modern Pakistan, shared Nehru's secularism even as he insisted on a separate Muslim state. Others in Pakistan called for this state to be subservient to the Islamic Shari'a. This division has brought the country to the brink of implosion over the never-healing sore of Kashmir and the recent reverberations of Islamic radicalism from Afghanistan and elsewhere in the Muslim world. There have been dark voices in Hinduism as well. The terrorism espoused by B. K. Tilak before World War I and the fascism of Subhas Chandra Bose in World War II found expression in the Hindu communalism of Sardar Vallabhbhai Patel, Nehru's co-prime minister in the first 2 years of independence. Patel died of a heart attack, but these several divisive strands collected into the Hindu nationalism of Mr. L. V. Advani and the Bharatiya Janata Party, which is now a coequal national rival to the secular Congress Party. India and Pakistan now confront each other as nuclear powers, and another *chakravartin*, in this tense subcontinent, is nowhere in sight (Mehta, 1996).

In China, the first coherent voice to articulate a path to modernization out of the swirling strands of the May Fourth Movement was Sun Yat-sen, who advocated *san min chuyi* (three people's principles) of people's livelihood, people's rule, and people's nationalism. The last was to uphold China's traditional Mandate of Heaven. The first was translated into rural life as "land to the tiller" (a theme that the communists later tried to call their own). For the second principle, democracy, Sun called for a transition to constitutional democracy in China through three stages of tutelage. In practice, Sun's political party, the Guomindang, could not pull it off. It lurched instead between the Christian social gospel of the New Life Movement and an Italian-like fascism of Blue Shirt discipline, all the while continuing in a reluctance to share power. Even as Sun's ideology failed in China, it became the basis for the subsequent economic miracle on Taiwan. It also describes the long path taken by South Korea to economic prosperity and a lagged following of this prosperity to full democracy (Wells, 2001).

Another failure was the Hu Shih liberals, who embraced linguistic reform and U.S.-style democracy. This faction was discredited by President Woodrow Wilson's treachery at the Treaty of Versailles (in acquiescing to granting the Concession of the Shandong Peninsula to Japan, rather than his public promise that it would be returned to China), even as it went on to discredit itself domestically by joining with the left-wing branch of the Guomindang in the strategic historical error of siding with the Japanese in their puppet state of Manchukuo.

The communists were the ultimate victors in both the civil war with the Guomindang and in the articulation of modern China. Although the form of government came straight from Lenin, Mao Zedong formulated a novel strategy of revolution—people's war—and introduced several innovative political projects and organizations, most of them disastrous. It was Deng Xiao Ping, the architect of China's unprecedented current economic growth, who reintroduced to China a pragmatism worthy of both Machiavelli and Adam Smith. This was reflected in his legendary question about the importance of the color of the cat as long as it could catch mice. The credit for this pragmatism, however, lay in the Four Modernizations of Deng's earlier protector, Zhou Enlai, who quietly made a career of fixing many of the excesses of Mao's zeal. It was an uneasy Gandhi–Nehru-like relationship, and China suffered for it—but might have suffered more without it (Goldman, 1994).

The truly novel definition of modernity in Asia came from Japan. Utterly defeated in World War II and under foreign occupation afterwards (1945–1952) for the first time in its history, Japan, in Article IX of its new constitution, outlawed war as an instrument of foreign policy and forbade the country to have anything but a minimal "Self-Defense Force" as a military institution. As a sovereign state, in what was called the Yoshida Doctrine, Japan placed its security in the hands of the United States and dedicated its own energies exclusively toward economic prosperity. Since then, in the era after the cold war, several intellectual and political voices have grown restive under this arrangement. One popular political writer, a former mayor of Tokyo, titled his recent book, *Just Say No—to the United States*. Others question the concept of nationality as an unwelcome Western transplant even as they articulate a distinctive identity and place for Japan (Sakai, de Bary, & Toshio, 2005).

Southeast Asia has continued to lament its strategic weakness. For nearly all Southeast Asian nations, modernization has been accompanied by outbursts of indigenous violence. It was convulsive in Indonesia in 1965 and again in 1998–1999. Burma, Thailand, Philippines, Malaya, Vietnam, Cambodia, and Laos all were wracked by insurgencies. Except for Malaya, the United States intervened in all of them, massively so in Vietnam. In these struggles, each country sought to define its own modern national identity in attempts to fashion integrative polities that

could overcome the separatist groups and ideologies fueling the insurgencies. With most of these convulsions over by the start of the new millennium (2000), these countries have now endeavored to integrate regionally. Their organization, the Association of Southeast Asian Nations, represents an interesting institutional counterpoise in international relations to the more developed European Union.

Conclusion: Cultural Grounding of Concepts

This consideration of the political thought of Asia as it has responded to the three contextual challenges of the classical, colonial, and modern periods brings us to the question of an Asian distinctiveness regarding modern Asian conceptions of democracy and its companion ideas of freedom and equality. Although the constitutions of many Asian states, those of India and Japan in particular, bear the imprint of Western ideas and institutions, the sources of these ideas emerge from different cultures and historical experiences, Asian ones. At root, although there is nothing in Asian experience or culture to preclude democracy itself, what may require different institutional expression of this principle is the fundamental difference between Asia and the West over the balance between the individual and the family. In all Asian countries, family and its ties to the state and its loyalties come before the freedom to chart individual destinies. In the West, on the other hand, individuals are encouraged to cut loose from family ties to freely chart their individual fortunes with no inequalities in status either within the family or in the larger society (at least in theory). This different balance calls for a different definitional relationship of freedom and equality to democracy. No one has made this distinction more clear than Lee Kwan Yew, the former prime minister of Singapore, who has insisted that democracy in Asia must still be subordinate to family discipline—and therefore made no apologies for authorizing the public caning of Western adolescents for vandalism in the streets of his city (Bell, 2000).

Hence, to discuss democracy in Asia, we need to bring other words and concepts into play. Really, democracy in Asia should be set in a discussion of statecraft and political authority. These issues, in Asia, were focused on creating order and preserving social hierarchy, although all Asian political systems recognized that statecraft and political authority were best served by reciprocity and the legitimating of their actions in ways that earned public support and approval. There are contextual grounds, then, for democracy in Asia, but not on the same egalitarian foundations as in the West. Pye (1985), for example, talks of democracy in Asia as best arising out of a historical context of paternal authority and what he calls a *politics of dependence*. Bell (2000) has proposed an Asian bicameral legislature, with one house based on popular egalitarian

representation and the other on knowledge, a “House of Scholars.” Parenthetically, this notion brings us around the intellectual circle to Plato’s insistence on ultimate rule by philosopher-kings (Lomperis, 1984).

Similarly, the Western centerpiece, freedom, needs to be recast in Asia as well. Rather than all the human rights guaranteed to individuals in the West through a constitutional Bill of Rights and the like, freedom in Asia has been differently defined in at least three ways. First, in Asia, freedom is more of a group concept than an individual one. Indians could pursue *swaraj* (self-rule) against the British, but to its greatest champion, Gandhi, for individuals *swaraj* meant more communal responsibilities to autonomous little communities (*ashrams*), not more individual human rights.

Thus, second, freedom for the individual boils down to relative degrees of autonomy from the multilayered obligations of these all-encompassing social structures. The overarching value here is responsibility. Freedom is the leftover. Daoist knights-errant and Hindu *kshatriya* warriors had the freedom of battle and of strategy, but only within the parameters of their larger duties to the Heavenly Mandate in China and the cosmic dharma (duties) of their souls in India. In the Indian epic *Mahabharata*, the hero Arjuna was not allowed the freedom to be a pacifist and opt out of the cosmic battle at Kurekshetra because the duty of his *kshatriya* caste compelled his martial service to uphold order. For women, duties were equally stark. In China, the virtues of high-class women were secured by footbinding. High-caste widows in ancient India had the “freedom” of avoiding the dejected status of widowhood or humiliating pollution of remarriage by committing *suttee* (self-immolation on a funeral pyre).

Third, the fullest expression of freedom in Asia is religious. In China, Buddhism offered release, or nirvana, from the world and its politics. Daoism cultivated a freedom of the soul within the external responsibilities and rituals of Confucianism. And in India, the householder (the responsible citizen, in Western parlance) could honorably flee to the forests, after discharging his many social and political duties, and seek *moksha*, the release that comes from enlightenment. Until the insertion of Western politics and ideas, freedom, in Asia, did not lie in politics.

Finally, the overarching Western ethos of equality has had a strong impact on all Asian societies. Indeed, this idea became the linchpin to undermining the Western imperium itself. But even with this wave of Western egalitarianism, Asian societies retain an even more profound rootedness in hierarchy. Western ideas of equal treatment and equal dignity have woven their way into the fabric of all Asian societies. But the “rightness” of hierarchy remains (Dumont, 1970). Gandhi, for example, called members of the “untouchable” caste harijans, or “children of God,” but still supported the moral virtue of the hierarchical caste system itself. Echoes of the old Confucian hierarchy

remain strong in China, as do patterns of the samurai ritual and hierarchical obligations in Japan, particularly in its unique corporate culture. Thus, equality in Asia, with this hierarchical persistence, is better rendered as *equity*, which is a word that gives more room for social ladders in a formulation of fairness and justice.

Illustratively, then, in passing these three universal political concepts of democracy, freedom, and equality through the analytical prism of the historical context of Asia, we find that all Asians persist in holding onto two anchors. First, Asians retain a strong preference for groups, particularly the extended family, over individuals as the primary unit of society. Second, in this preference for groups, Asians continue to choose a hierarchical ordering of these groups over any comprehensive notions of full social equality. The persistent hold of these two anchors necessitates an Asian reformulation of these concepts, which have heretofore been defined only from a Western context. Thus, the expression of individual freedom or rights from the West must in Asia be tempered by a greater consideration for group responsibilities so that freedom in Asia is merely a relative autonomy from them. Similarly, the penchant for hierarchy in Asia imposes equity as an appropriate expression of fairness, rather than equality. Turning to politics from these two reformulations, democracy in Asia, therefore, will need to be constructed and expressed in political arrangements that value groups and legitimate hierarchy. Thus, the cultural settings of such seemingly universal political concepts as democracy, freedom, and equality achieve richer meaning and nuance when analyzed comparatively through their evolution in other cultures, including those in Asia.

References and Further Readings

-
- Basham, A. J. (1959). *The wonder that was India*. New York: Grove Press.
- Bell, D. A. (2000). *East meets West*. Princeton, NJ: Princeton University Press.
- Benedict, R. (1946). *The chrysanthemum and the sword: Patterns of Japanese culture*. Boston: Houghton Mifflin.
- Brown, D. M. (1954). *The white umbrella: Indian political thought from Manu to Gandhi*. Berkeley: University of California Press.
- Brown, D. M. (1965). *The nationalist movement: Indian political thought from Ranade to Bhave*. Berkeley: University of California Press.
- Clarke, J. J. (1997). *Oriental enlightenment: The encounter between Asian and Western thought*. New York: Routledge.
- Dallmayr, F. R. (Ed.). (1996). *Beyond orientalism: Essays on cross cultural encounter*. Albany: SUNY Press.
- DeBary, W. T. (1983). *The liberal tradition in China*. New York: Columbia University Press.
- DeBary, W. T., & Weiming, T. (Eds.). (1999). *Confucianism and human rights*. New York: Columbia University Press.
- Dumont, L. (1970). *Homo hierarchicus: The caste system and its implications*. Chicago: University of Chicago Press.

- Gandhi, M. K. (1957). *An autobiography: The story of my experiments with truth*. Boston: Beacon Press.
- Gluck, C. (1985). *Japan's modern myths: Ideology in the late Meiji period*. Princeton, NJ: Princeton University Press.
- Goldman, M. (1994). *Sowing the seeds of democracy in China: Political reform in the Deng Xiaoping era*. Cambridge, MA: Harvard University Press.
- Hall, D. L., & Ames, R. T. (1999). *The democracy of the dead: Dewey, Confucius, and the hope for democracy in China*. Lasalle, IL: Open Court.
- Hardgrave, R. L., & Kochanek, S. A. (2000). *India: Government and politics in a developing nation*. Fort Worth, TX: Harcourt Brace.
- Jansen, M. B. (1995). *Japan and its world*. Princeton, NJ: Princeton University Press. (Original work published 1975)
- Johnston, A. (1998). *Cultural realism: Strategic culture and grand strategy in Chinese history*. Princeton, NJ: Princeton University Press.
- Kelley, D., & Reid, A. (Eds.). (1998). *Asian freedoms: The idea of freedom: East and Southeast Asia*. Cambridge, UK: Cambridge University Press.
- Kershaw, R. (2001). *Monarchy in South East Asia: The faces of tradition in transition*. New York: Routledge.
- Levenson, J. R. (1958-1968). *Confucian China and its modern fate* (Vols. 1-3). Berkeley: University of California Press.
- Lomperis, T. J. (1984). *Hindu influence on Greek philosophy: The odyssey of the soul from the Upanishads to Plato*. Calcutta, India: Minerva. (Republished on the Internet by IdeaIndia.Com, 2007.)
- Malik, H. (Ed.). (1971). *Iqbal: Poet philosopher of Pakistan*. New York: Columbia University Press.
- Marshall, S. J. (2002). *The mandate of heaven: Hidden history in the I Ching*. New York: Columbia University Press.
- Mehta, V. R. (1996). *Foundations of Indian political thought: An interpretation (from Manu to the present day)*. Columbia, MO: South Asia Books.
- Metcalf, T. R. (2001). *Ideologies of the Raj*. Cambridge, UK: Cambridge University Press.
- Muller, M. (1919). *India: What can it teach us?* London: Longmans, Green.
- Nehru, J. (1959). *The discovery of India*. Garden City, NY: Doubleday. (Original work published 1946)
- Northrop, F. S. C. (1946). *The meeting of East and West, an inquiry concerning world understanding*. New York: Macmillan.
- Ogden, S. (2002). *Inklings of democracy in China*. Cambridge, MA: Harvard University Press.
- Perry, E. J. (2001). *Challenging the mandate of heaven*. Armonk, NY: M. E. Sharpe.
- Pye, L. W. (1985). *Asian power and politics*. Cambridge, MA: Belknap.
- Radhakrishnan, S., & Moore, C. A. (Eds.). (1957). *A sourcebook in Indian philosophy*. Princeton, NJ: Princeton University Press.
- Sakai, N., de Bary, B., & Toshio, I. (Eds.). (2005). *Deconstructing nationality*. Ithaca, NY: Cornell University Press.
- Schwartz, B. I. (1985). *The world of thought in ancient China*. Cambridge, MA: Harvard University Press.
- Seven military classics of ancient China, *The*. (1993). (R. D. Sawyer, Trans., with M. c. L. Sawyer). Boulder, CO: Westview Press.
- Spear, P. (1961). *India: A modern history*. Ann Arbor: University of Michigan Press.
- Tambiah, S. J. (1976). *World conqueror and world renouncer: A study of Buddhism and polity in Thailand*. Cambridge, UK: Cambridge University Press.
- Tucker, R. C. (Ed.). (1972). *The Marx Engels reader*. New York: W. W. Norton.
- Wells, A. (2001). *The political thought of Sun Yat Sen*. New York: Palgrave.
- Wittfogel, K. A. (1957). *Oriental despotism*. New Haven, CT: Yale University Press.
- Yuzan, D. (1941). *The code of the samurai* (A. L. Sadler, Trans.). Rutland, VT: Charles E. Tuttle.
- Zhou, Cezong. (1960). *The May Fourth movement: Intellectual revolution in modern China*. Stanford, CA: Stanford University Press.

67

ISLAMIC POLITICAL THOUGHT

SVETLA BEN-ITZHAK

University of Kansas

Islamic political thought has found numerous expressions from its very beginnings through modern times. As such, presenting an overview of its development and essence is not an easy task as one must include not only the contributions of key Islamic political philosophers (e.g., al-Farabi, Ibn-Khaldoun, and Avicenna), but also the religio-political credos of main Islamic schools of thought (i.e., the Sunni, the Shi'a, and the Kharijites), as well as an overview of main Islamic political concepts (i.e., the question of succession and leadership, or the *khalifah* versus the *imamah* precepts). Although some have argued that Islamic political thought has been neglected by all but a few specialists, several scholars have undertaken the arduous task of exploring, discussing, and summarizing the milieu, meaning, and significance of Islamic political thinkers, ideas, and developments (see Ayoob, 2007; Black, 2001; Crone, 2005; Enayat, 1982; Lewis, 1991).

Through philosophers, concepts, and religious movements, Islamic political thought has impacted religious, secular, and academic communities. Thus, given its intellectual and practical significance, the need to more fully comprehend the dynamics and development of Islamic political thought warrants and justifies the inclusion of this chapter in this handbook. First, this chapter defines the concept of Islamic political thought in terms of its etymology. Second, it surveys the classical period of Islamic political thought, including its origins, the religious-political schism within Islam, and the contributions of key classical

Islamic political philosophers. Third, the chapter summarizes modern expressions of Islamic political thought and current academic research on the topic. The conclusion recaps the main points in the historical and academic development of Islamic political thought.

Etymology of Islamic Political Thought

As this chapter views philosophy and thought as interchangeable concepts, it defines the latter in terms of the former. Generally, political philosophy refers to the study of state affairs and processes as well as to the in-depth search for rationales in politics and ethics in public behavior (Walzer, 1963). To this, Islamic political thought adds a specific framework—that of Islam—to study, explain, and rationalize all things political. This framework is derived from the very sources of Islam: the Qur'an (comprising revelations to Muhammad), supplemented by the *hadiths* (stories about Muhammad's life, words, and deeds). Thus, Islamic political thought began from the very inception of Islam (circa 622), and its development is generally divided in two main periods: *classical or premodern* (645–1500), from the historical origin of Islam to the end of the classical period, and *early modern and modern* Islamic political thought (1500–present), which includes the dynastic period from the Safavid empire to contemporary Islamic political movements.

Premodern Islamic Political Thought

It would be impossible to include all thinkers, concepts, and movements that are part of this period of Islamic political thought. Such an endeavor should begin with Muhammad's life and his political contributions to the first Islamic state, the religious and political schism within Islam, a discussion of the Sunni tradition (*sunna*), and the Shi'ite theories of leadership (*imamah*; 622–1000), an overview of the theory of the caliphate in *din wa dawla* (religion and the state; 1000–1220); and an explanation of Shari'a ideology and the spread of Islam (1220–1500). In addition, a summary of the works of major Islamic political philosophers and thinkers during this period should be included.¹ All this is summarized briefly next beginning with Muhammad's life as both a religious and political leader.

Muhammad as a Political and Religious Leader

Muhammad's position in the early Muslim community was of God's appointed religious, political, and military leader, which was a central factor in keeping the Muslim community united both religiously and politically. Thus, Muhammad's death in 632 CE suddenly posed several questions to his followers: Who should succeed Muhammad? What is the Islamic way of choosing his successor? What type of government should the Muslim community have? The answers to these questions were further complicated as Muhammad died without giving clear instructions as to how to choose a successor or the type of government that the community needed to establish.

As a result, the death of Muhammad in 632 CE started a disagreement over who should succeed him as a political and religious leader of the *umma*—the Muslim community. While Muhammad's body was being prepared for burial by his close family (including Ali Ibn Abi Taleb, Muhammad's cousin and son-in-law), a small group of the *Ansar* (Muhammad's helpers) met in Saqifat Bani Sa'eda and selected Abu Bakr, a close companion of Muhammad, as his successor (and first caliph). However, some did not accept this decision, arguing that Ali (Ibn Abi Taleb) was appointed as Muhammad's successor by Muhammad himself (Lapidus, 2002).

Since that time, succession (*khalifah* or *imamah*, caliphate system or imamate system), and how it should be done, has remained a matter of dispute. The Umayyads—the second Islamic caliphate (after the Rashidun caliphate of the first four caliphs) established after the death of Mohammed—tried to resolve the issue by decreeing that the caliph should appoint his own successor. Thus, before Abu Bakr died, he chose Umar as his successor. Umar, on the other hand, delegated the election of his successor to the Majlis al-Shoura, a consultative council (Subhi, 1969). This council chose Uthman as the third caliph, but a rebellion against him instituted Ali as the fourth caliph and led

to the first civil war among the Muslims in the Battle of *Al Jamal* (the camel). Following Ali's succession, a second major civil war, the Battle of Siffin, ensued between Ali's supporters and the supporters of Mu'awiyah Ibn Abi Sufian, who founded the Umayyad dynasty (Ismael & Ismael, 1980; Subhi, 1969).

In summary, early Islamic religious and political history reflects the following: First, Muhammad did not indicate how to choose a successor, nor did he indicate a preference for one form of state over another (nor did the Qur'an). Second, in the few years following the death of Muhammad, four distinct patterns of succession emerged: (1) limited choice (Abu Bakr), (2) nomination by the caliph (Umar), (3) *shoura*, or consultation (Uthman), and (4) a coup (Ali). Third, the question of succession led to a greater political conflict that, after the assassination of Uthman, caused a religious schism in Islam, dividing the *umma* into Sunni and Shi'a. In addition, the establishment of the Umayyad dynasty and the violent death of Ali's son Hussein further affirmed and deepened the gap between the Sunnis and the Shi'as, consolidating their differences in religious and political matters (Ismael & Ismael, 1980; Subhi, 1969).

Furthermore, the dispute over the question of succession resulted in a distinction between the caliph and the imam. Originally, *khalifah* was used to denote a caliph's religious and political leadership (the caliph is also known as *Amir al Mu'minin*, or Commander of the Believers). The word *imam*, on the other hand, referred to the leader of the Muslims in prayers. However, with the establishment of the Umayyad dynasty, supporters of Ali, the Shi'a, developed the doctrine of the *imamah*, which gave both religious and political dimensions to the ruler. The Shi'a, therefore, upheld the view that the Imamate belongs to Mohammad and his descendants, of which Ali was the first Imam. In addition, Shi'a doctrine affirmed that the Imam is incapable of sin (Hairi, 1977). The Sunnis, on the other hand, accepted the caliph as a temporal ruler who is "pre-eminently a political functionary and though he may perform religious functions, these functions do not imply the possession of any spiritual powers setting him thereby apart from the rest of the faithful" (Arnold, 1965, p. 17). Thus, in Sunni tradition, the caliph did not enjoy any particular religious role although Sunnis continue to see the Imamate, as reflected in the rules of the first four successors, to be the true Islamic form of government.

The conflict over succession led to the creation of three main schools of Islamic political thought: Sunni, Shi'a, and Kharijites. These are briefly discussed next.

Three Schools of Islamic Political Thought

For both Sunnis and Shi'a, the imam became "the de jure ruler and the de facto caliph" (Subhi, 1969, p. 24). However, early Muslims were confronted with myriad

questions as to the nature and extent of the imam's power. How are imams chosen? What is their role? What is their relationship with the rest of the Muslim community? These questions bridging political reality and religious roles were never truly answered, at least not in a unanimous way.

The Sunni School

The Sunnis saw the role of the imam as essential for the well-being of the community, but they did not bestow on him any mystical powers (Hairi, 1977). To choose the imam, the Sunnis emphasized the role of *Ahl al Hal wa al 'Aqd* (those who loose and tie, meaning those who know; a select few) to legitimize the first four caliphs. Under this doctrine, the Muslim community was not given an active say in choosing its ruler, which eventually led to dissension, factionalism, and rebellions. However, instead of discussing the issues caused by the *Ahl al-Hal wa al-'Aqd* doctrine, Sunni theorists focused on other problems, such as how many people should belong to *Ahl al-Hal wa al-'Aqd* and how many are needed to have a legal *bay'ah* (pledge of allegiance). Muhammad al-Baqelani (died 1013), for instance, established several rules in that regard: The nomination of the imam (*'aqd al imama*) can be done by one person, and six people are needed to achieve *bay'ah*. In addition, the community cannot depose an imam unless he denounces Islam or stops praying and teaches others to do the same, the imam is not infallible, he has to come from the tribe of Quraysh, and he needs to be knowledgeable about war and how to protect the community (Ibish, 1966; Ismael & Ismael, 1980).

Some Sunni religious thinkers were challenged in their views by Muslim philosophers who were reading and commenting on Greek philosophers such as Plato and Aristotle. One of these philosophers was al-Farabi (870–950) who, in his book *Ahl al Madinah al Fadelah*, introduced his theory of the state. Al-Farabi affirmed that the state should be based on a “mutual renunciation of rights,” for in a society or a state, each individual must give up “in favor of the other a part of that by which he would have overpowered him, each making it a condition that they would keep perfect peace with each other and not take away from the other anything except on certain conditions” (quoted in Sherwani, 1977, p. 71). Also, al-Farabi defended Plato's philosopher king, describing his *ra'is awwal* (ideal ruler) as one “who by his very nature and upbringing, does not want to be instructed by others and who has the inherent capacity for observation and of conveying his sense to others” (quoted in Sherwani, 1977, p. 73). For the perfect *ra'is awwal*, al-Farabi enumerated 12 traits but was satisfied with one who had only 5 or 6. In addition, if no man was found possessing even that minimum number of attributes, then a council with two to five members having an aggregate of 10 attributes among them should be chosen. One

of these two to five council members, al-Farabi asserts, must be a *hakim*—a wise man who knows the needs of the state and its people.

Another prominent Sunni theorist, Abu al-Hassan al-Mawardi (975–1058), also wrote on the role of the imamah. In his works *Ahkam Sultaniyah*, *Nasihah al Muluk*, and *Qawanin al Wizarah*, al Mawardi argues that the imamah is a “caliphate of prophethood in safeguarding religious and temporal affairs” (quoted in Sherwani, 1977, p. 148) and, as such should follow the “right path,” guided by both the Shari'a and reason. The imam, on the other hand, should be just, knowledgeable of the purposes of *ijtihad* (i.e., making a legal decision based on an independent interpretation of the Qur'an and the sunna), without physical or sensual handicaps, wise in ruling the community and its affairs, courageous in protecting Islam, and a descendant of the Quraysh tribe (Ismael & Ismael, 1980). His role is to protect the faith, judge among people, punish transgressors, appoint just men, and lead a life between piety and luxury. As to how the imam should be chosen, al-Mawardi prescribes either by *Ahl al-Hal wa al-'Aqd* or election or by the previous imam. The imam, argues al-Mawardi, can be removed for only two reasons: *hajr* (acting against Islam) and *qahr* (falling prisoner and having no hope of being saved or freed; Ibish, 1966; Sherwani, 1977).

A later Sunni religious thinker, Ibn Taimiyah (1263–1328), who wrote *Minhaj al Sunnah*, also did not consider it necessary for the community to elect its imam, but he affirmed that the *bay'ah* (the pledge of allegiance) served as a bilateral contract between the Imam and his community. Although, to Ibn Taimiyah, the imam could reach a certain point of perfection, he could never be a prophet. In contrast to other Sunni thinkers, however, Ibn Taimiyah did not affirm that the Imam should be from the Quraysh tribe as he saw this condition contradictory to Islamic teachings of equality among Muslims. Yet he also argued that although the rule of the imam is vested in divine law, the imam should consult *Majlis al-Shoura*, or the consultative council, and be guided by its decisions more than by historic precedents (Sherwani, 1977).

In summary, Sunni political thought generally did not perceive the nature or origin of political power as problematic and focused instead on its implementation and administration. As such, the main issues addressed by Sunni theory were the personal qualities of the ruler and the organization of the state “since neither the attainment of power nor its fundamental legitimacy were at issue for the Sunni sect” (Ismael & Ismael, 1980, p. 605).

The Shi'a a School

The religious doctrine of Shi'a evolved as a political protest regarding the disputed question of succession following Muhammad's death (Subhi, 1969). However, although the conflict began immediately after Muhammad's demise, it did not clearly delineate specific religious

ideologies until after the death of the fourth caliph, Uthman and the ensuing controversy between the supporters of Ali and the Umayyad dynasty (Ismael & Ismael, 1980). A central tenet of the Shi'a doctrine was and is that the imamah belonged only to Muhammad and his descendants. Thus, the imam, like Muhammad, is believed by the Shi'a to be an infallible interpreter and protector of the law (Sherwani, 1977; see also Ismael & Ismael, 1980). A good description of the Imam's powers is provided by al-Imam al-Rida in Subhi (1969).

The largest branch of the Shi'a are the Twelvers. They believe that there were 12 divinely ordained imams who are the spiritual and political successors of Muhammad, with the first rightful imam being Ali (the Islamili and the Zaydi, while also Shi'a, believe in a different number of imams). Each succeeding imam was the son of the previous imam (with the exception of Husayn ibn Ali, who was the brother of Hasan ibn Ali). The 12th and final imam was Muhammad al-Mahdi, who is believed by the Twelvers to be currently alive and in hiding.

Thus, generally, Shi'a religious and political thought was governed by the doctrine of the occultation of the last imam, his reappearance, and the hope for his rule in a just and egalitarian manner in accordance with Islamic laws and precepts. In the expectation of the Imam's return, Shi'a thinkers developed the doctrine of *taqiyah*—religious and political beliefs that have been used to justify an acceptance of existing governments. In addition, this doctrine has stimulated research on the nature of political authority during the period of the greater occultation of the last imam. In general, the contradiction between the ideal state led by the last imam and the necessity of government in the meantime has led Shi'a intellectuals to elaborate and discuss the structure and functions of political power in the less-than-ideal state. As such, concepts of constitutionalism and democracy have been integrated into Shi'a concepts of government during this period of absence. Thus, Shi'a political thought has focused mainly on the nature and origins of power during the imam's absence, the limitations of the usurpation of power, and the accountability of the leaders. It is argued that Ayatollah Khomeini's political thought is a direct continuation of this trend of Shi'a religious and political theory (Ismael & Ismael, 1980).

In addition, following the great occultation of the last imam, Muhammad al-Mahdi, the doctrines of *marja i taqlid* (a source to follow or a religious reference) and *ijtihad* became important and were used by religious scholars or the *ulama* to provide leadership for people in religious, social, and political matters (Hairi, 1977). As such, after the Qur'an and the imams, *marjas* have the highest religious power in Twelver Shi'a Islam. As a result, a Grand Ayatollah (title granted to a top Shi'a *mujtahid*, or a male Shi'a scholar competent to interpret Shari'a using *ijtihad*, or independent thought, has the authority to make political and religious decisions for the *umma*).

One of the influential religious leaders in Iran at the beginning of past century was Mirza Muhammad Husayn Na'ini, who tried to reconcile the ideas of constitutionalism (*mashruti*) with the principles of Islamic government. To do this, he established two general purposes of a government, namely, (1) to maintain internal order, protect individual rights, and support education and (2) to prevent any foreign intervention by "preparing a defensive force and war ammunitions, and the like" (Hairi, 1977, p. 166). Na'ini asserted that there are two kinds of governments: *tamallukiyyah* (tyranny) and *vilayatiyyah* (constitutional government). He argued that Muslims need to fight *tamallukiyyah*, for it turns human beings into slaves. On the other hand, while Na'ini held to the Shi'a doctrine that the ideal government is the government of the imam who is infallible, sinless, and possesses God-given knowledge, in the absence of such an imam, "the only possibility left . . . is to choose a constitutional form of government, even though the latter would still be a usurpation of the Imam's authority" (Hairi, 1977, p. 191). However, the ruler in such government must still gain the approval of the *ulama*. Hence, Article 2 of the Iranian Constitution includes the following provision:

At no time must any legal enactment of the National Consultative Assembly . . . be at variance with the sacred principles of Islam . . . It is hereby declared that it is for the . . . *ulama* . . . to determine whether such laws as may be proposed are or are not conformable to the principles of Islam; and it is therefore officially enacted that there shall at all times exist a committee of not less than five *mujtahids* . . . so that they may . . . reject and repudiate, wholly or in part, any such proposal which is at variance with the Sacred Law of Islam, so that it shall not obtain the title of legality. (Hairi, 1977, p. 213)

This article clearly gave the religious authorities an official role in determining the compatibility of political laws with Islamic principles. However, some argue that this article, like the constitution itself, remained abstract as political leaders effectively counteracted these limitations of their powers (Ismael & Ismael, 1980).

The Kharijites School

The Kharijites (renegades) school of thought maintained that the leadership of the Muslim community should be open to all Muslims and that an elected caliph should not relinquish his right under any circumstances. However, if he is unjust, he should be deposed by any means (Hassan, 1967). In fact, the Kharijites argued that the caliphate and imamah were not necessary. One of the Kharijites' main thinkers, Shahrastani, affirmed that the "Imamah is not necessary according to *Shari'a* (Islamic law). It is based on people's interactions with each other. If everyone justly deals and cooperates with the others, as well as fulfills

his duties and responsibilities, they do not need an Imam” (quoted in Subhi, 1969, p. 69; see also Hassan, 1967, p. 161).

Thus, this school of thought introduced a rather radical view of Islamic political structure. Many of its theories became popular in the 19th century with the advent of nationalism. As such, some Islamic reformers, including Jamal al-Din al-Afghani (1839–1897) and Mohammad Abdo (1848–1905), tried to reconcile Western democratic precepts with the Islamic idea of the state. The combination of the Kharijites’ theories and the teachings of such thinkers led to a new generation of secular thinkers who wanted to completely separate religion from the state. Ali Abd al-Razeq (1972) is one of the best representatives of these new ideas. Writing in 1925, he maintained that Islam never promoted or decided on a particular form of government, nor were Muslims ever required to build a particular system. Abd al-Razeq did not perceive the caliphate to be a religious system and argued that the Qur’an did not order it. Al-Razeq (1972) disentangled Islam from the caliphate and argued that “Islam is innocent from the caliphate because it paralyzed any development in the form of government among the Moslems” (pp. 92–93). More basically, al-Razeq asserted that Muhammad did not form a government or establish a state, and thus the caliphate in Islam was “based on nothing but brutal force,” a catastrophe that “hit Islam and Moslems and is a course of evil and corruption” (pp. 129, 136). Al-Razeq, therefore, maintained that Muslims are free to choose their form of government (Ismael & Ismael, 1980).

In addition to this religio-political schism in early Islam, several Islamic philosophers also contributed to shaping early Islamic political thought and are discussed next.

Classical Islamic Political Philosophers

As political philosophy is, in general, concerned with the search for understanding what constitutes a good or the best political regime, a Muslim political philosopher would similarly contemplate political structures and regimes. However, he must temper his conclusions with what the Islamic tenets hold to be a good or the best political regime. The contributions of key classical Islamic philosophers including al-Farabi (Abunaser), Ibn Sina (Avicenna), Ibn Bajjah (Avempace), Ibn Rushd (Averroes), and Ibn Khaldun are briefly presented next.

Al Farabi

Abu Nasr Muhammad al-Farabi (circa 872–950 or 951), also known in the West as Alfaribius, was a Muslim philosopher, scientist, and musician.² Walzer (1963) calls al-Farabi and Ibn Khaldun “the most outstanding political thinkers in the Islamic world” (p. 43) and argues that al-Farabi was ultimately convinced of the primacy of human

reason and philosophical truth. However, as there are very few extant reliable sources (i.e., composed prior to 12th century) much is unknown of al-Farabi’s actual life and work (Mahdi, 2001).

Generally, al-Farabi’s political writings discuss the nature of the virtuous regime whose guiding principles he affirmed to be the realization of human excellence by the promotion of virtue. He believed that man is not rendered perfect simply through the natural principles found in him but can achieve perfection through rational understanding and deliberation. The ultimate perfection a man can achieve is by attaining the supreme happiness available to him. “Happiness is the good desired for itself, it is never desired to achieve by it something else, and there is nothing greater beyond it that a human being can achieve” (Dieterici, 1895, p. 46; cf. *Ketab al stasa*, pp. 72–75, 78). Thus, al-Farabi saw the ideal society as one promoting and seeking the realization of “true happiness” (Butterworth, 2005). However, al-Farabi thought that such “virtuous” societies are rare and, with most falling short of attaining this goal, he divided them into three categories: ignorant, wicked, and errant (democratic societies fall into the category of “ignorant” as he perceived them as lacking guiding principles).

As to the role of religion, al-Farabi affirmed that the ancients saw religion as “an imitation of philosophy. Both comprise the same subjects and both give an account of the ultimate principles of the beings. . . . In everything in which philosophy gives an account based on intellectual perception or conception, religion gives an account based on imagination. In everything demonstrated by philosophy, religion employs persuasion” (quoted in Lerner & Mahdi, 1963, p. 77). As such, al-Farabi perceived that the imam, the philosopher, and the legislator were a single idea. The implication thereof is that Muhammad was the ultimate philosopher-king and that philosophers are superior to those who are merely religious. Thus, for al-Farabi, the real division was between those who can ground their beliefs philosophically and those who cannot (termed *simple believers*).

Academics argue as to whether al-Farabi actually outlined any political program in his writings. Corbin (1993) affirms that al-Farabi’s philosophical contributions need to be considered “prophetic philosophy” rather than political works. Similar to Corbin, Reisman (2005) argues that al-Farabi did not advance any political doctrine per se but depicted different types of societies, not for political aims, but as a part of ethical discussion. Butterworth (2005), on the other hand, asserts that al-Farabi targeted mainly the “king” and “statesmen,” and nowhere in his works does he speak of the “prophet-legislator” or of revelation. Finally, Parens (2006) argues that al-Farabi shrewdly mixed Islam and philosophy by demonstrating that too many conditions need to be met in order for a pan-Islamic virtuous society to be achieved.

Ibn Sina

Avicenna (full name: Hussain ibn Abdullah ibn Hassan ibn Ali ibn Sina; circa 980–1037), also known in Persian as Abu Ali Sina or in Arabic as Ibn Sina, was a Persian scientist, logician, poet, and philosopher (Mahdi et al., 2008). Avicenna was born in approximately 980 in present-day Uzbekistan, near Bukhara, and died in 1037 in the Iranian city of Hamedan. He wrote some 450 treatises on a variety of topics, but only 240 have survived, among which 150 discuss philosophy. Avicenna's most famous works are *The Book of Healing*, which is a philosophical and scientific encyclopedia, and *The Canon of Medicine*, a medical textbook. Avicenna pioneered the so-called Avicennian logic, as well as the philosophical school of Avicennism, both of which have remained influential among later Muslim thinkers. Although Avicenna is viewed mostly as the father of modern medicine and clinical pharmacology, he also wrote on early Islamic philosophy, including logic and ethics. Avicenna outlined philosophical reasons for the practice of religion and viewed the purpose of a caliph as a successor of Muhammad while the imam was a religious leader of the umma (Butterworth, 1992).

Ibn Bajjah and Ibn Tufayl

Abu-Bakr Muhammad ibn Yahya ibn al-Sayigh (1095–1138), also known as Ibn Bajjah, was a Muslim astronomer, logician, philosopher, musician–poet, and scientist. In the West, he was known as Avempace. He was from Andalusia, originally born in Zaragoza (today's Spain) but died in Fez, Morocco (McGinnis, 2007). In contrast to al-Farabi's political Platonism, which identified Muhammad with a philosopher king ruling a state founded on and governed by a divine law, Ibn Bajjah (and subsequently his student Ibn Tufayl) did not mix philosophical thought with practical politics. As Leaman (1980) writes, "Ibn Bajja's problem is how the philosopher in the imperfect state should relate to society" (p. 110) or, in other words, both Ibn Bajja and Ibn Tufayl wrestled with the notion of "the imperfect state" and the role of the Islamic philosopher in it. This is also the central theme in Ibn Bajjah's *Tadbir al mutawahhid* (The Governance of the Solitary) and Ibn Tufayl's *Hayy ibn Yaqzn* (Son of Awake—also known as Philosophus Autodidactus).

Ibn Rushd

Averroes (circa 1126–1198), or Abu'al-Walid Muhammad ibn Ahmad ibn Rushd and also known as Ibn Rushd, was a scholar of Islamic philosophy, theology, politics, music, and science. While Al-Farabi and Avicenna lived in the eastern part of the Islamic world, Averroes lived in Spain, in the city of Cordova. There he wrote several commentaries on Aristotle, but being a prolific writer, Averroes also covered numerous subjects,

including early Islamic philosophy, logic, jurisprudence, politics, and medicine.

Although Averroes mostly discussed other topics, in his commentaries on most of Aristotle's logical and theoretical work and on Plato's *The Republic*, Averroes argued that the ideal state, as described by Plato, has in fact the same constitution as the original Arab caliphate (Ahmad, 1994). Furthermore, similar to other Islamic political philosophers who used treatises and commentaries on rhetoric to express political and philosophical thoughts, Averroes wrote two commentaries on Aristotelian rhetoric, namely, *Talkhis* and *Summary*. In them, he not only developed ideas on political philosophy but also instructed others how to conduct political investigation.

Similar to al-Farabi, Averroes shared the view that Islamic philosophical thought has political implications while examining the linkage between law and philosophy, in his *Fasl al maqal* (Decisive Treatise on the Connection Between Religion and Philosophy). In it, Averroes affirmed that the law commands the study of philosophy for those who are capable of it. In other words, philosophizing is a duty that is not an idle enterprise but a necessary one as the law revealed by Muhammad is a divine law given for the benefit of the entire community. Thus, those who are capable of philosophizing (i.e., the philosophers) are obliged to use their wisdom to explain the law for the benefit of all. In arguing so, Averroes links practical political philosophy to Islam, thereby conceiving of a society grounded in obedience to a divine law as interpreted through people appointed to that end. Thus, similar to Al-Farabi, Averroes envisioned a state in which the philosophers are the elite and the leaders, and he argued that philosophy and Islam are in harmony and not in conflict (Butterworth, 1972).

Ibn Khaldun

Abu Zayd 'Abdu r-Rahman bin Muhammad bin Khaldun Al-Hadrami, or Ibn Khaldun (circa 1332–1406), was an Islamic theologian, historian, scientist, and philosopher. He was born in present-day Tunisia and became best known for his *Muqaddimah* (or *Prolegomenon*), which was the first volume of *Kitab al Ibar*, a book on universal history in which he attempted to decipher whether historical sources are credible. As such, living in the 14th century, Ibn Khaldun dedicated himself to the historical task of analyzing the rise and decline of state power in the Islamic world. His main subject was the political reality of the Islamic world, with the caliphate as a union of political and spiritual authority (Walzer, 1963). In the *Muqaddimah*, a lengthy treatise on political and social processes of change, Ibn Khaldun asserted that no individual society or dynasty can remain at a permanently high level of development, for once a maturity is attained, decay begins. In addition, Ibn Khaldun argued that, in the social and political development of a society, a sense of solidarity or a group feeling is

important, for it makes individuals identify with the group and subordinate personal interests to state interests. Among the natural bases for such solidarity to develop (i.e., kinship), ibn Khaldun perceived religion to be a powerful catalyst fostering a strong group belonging.

Summary

Islamic philosophical and theoretical inquiry into the role and nature of government, as well as its relationship to religious affairs and concepts, has been evolving for 1,300 years. Some argue that Western scholars have given only scant attention to understanding the theoretical and intellectual basis of Islamic governments, perceiving Islam instead as a “traditional” force that has resisted progressive change and has been a barrier to social and political development (Ismael & Ismael, 1980). Those scholars also affirm that this failure to recognize the ongoing theoretical and intellectual ferment in Islamic thought faced by the emergence of unique contemporary issues within Islamic societies “has contributed to the image of Islam as a static rather than dynamic system of thought incapable of relevance to the complex modern Issues of social and political development” (Ismael & Ismael, 1980, p. 601). Islamic political thought has not only undergone change in response to new conditions but continues to do so, as the next section outlines.³

Modern Islamic Political Thought

To understand the development of Islamic political thought during this period would mean to understand the historical background of modern Islamic Arabia, along with the political, social, economic, religious, and geographic situations that have shaped the dynastic, colonial, and contemporary eras experienced by Islamic societies. It would also mean understanding the main political movements of modern Islam, namely, the dynastic and modern-day caliphate, pan-Islamism, and pan-Arabism, as well as myriad Islamic political trends. It would also mean learning more about modern issues faced by current Islamic societies, including those posed by imperialism, neocolonialism, debates over religious and political authority in Islamic states, and the rise of violent resistance. It is, therefore, beyond the scope of this chapter to summarize everything that has impacted postclassical Islamic political thought, but a number of studies have already studied this topic at length (see Aslan, 2006; Black, 2001; Enayat, 1982; Lapidus, 1996).

However, generally, modern Islamic political thought has found expression through several often contradictory movements. The first is *traditionalist fundamentalism*, which endorses the traditional commentaries on the Qur'an and the sunna, making *taqlid* or imitation its basic

mode of operation and, thus, refusing to innovate. Generally, these movements follow one of several Islamic legal schools or *madh'hab*. The four main *madh'hab* of Sunni jurisprudence today (named after their founders *A'immah Arba'a*, or the four imams of *fiqh* (Islamic jurisprudence) are the *Hanafi*, the *Maliki*, the *Shafi'i*, and the *Hanbali* schools of law. The Shi'a legal school is the Jafari *madh'hab*, named after the person believed to be the sixth imam (Roy, 1998). The second stream of current Muslim political thought is *reformist fundamentalism*, which usually develops as a response to a perceived external threat and aims to return to the founding, original texts (Roy, 1998). Some examples would include the political-religious movements begun by Shah Wali Allah in India, Abd al-Wahhab in Arabia, and the Salafite faction. The third is *Islamism* (also called *political Islam*), which calls for a return to Shari'a law and often uses extreme means to achieve its goals. Modern-day examples include the Jamaat-e-Islami, Muslim Brotherhood, and the Iranian Islamic Revolution (Halliday, 1995; Tibi, 1998). A fourth strain is *liberal movements within Islam*, which generally oppose the first three Islamic political movements yet often share their anti-Western, anti-imperialist elements.

In the past, Islamic political movements formed in response to colonialism and Western imperial expansions. Today, debates and divisions in Islamic political thought gravitate around several core issues, including, among others, Zionism and the response to the formation of the state of Israel, the status of women and women's rights, and Islamic economies and the debt of Muslim states. Some Islamic activists also blame the maladies of Muslim nations on the influx of foreign or Western ideas and practices, including the spread of capitalism, feminism, and purported Muslim persecution by the West.

Thus, current academic research on Islamic political thought has generally focused on several main themes. In light of the recent Islamic revolution in Iran, some scholars have (re)examined the ideas of leading modern Muslim thinkers on the relationship of Islam with government (see Black, 2001; Enayat, 1982; Sivan, 1985). Others have chosen to explore a recurrent theme in a growing amount of research, namely, the relationship between Islam and democracy (see Diamond, Plattner, & Brumberg, 2003; el Fadl, Cohen, & Chasman, 2004; Esposito & Voll, 1996; Tamimi, 1997). Scholars have also investigated the linguistic dissimilarities between Islamic and Western political thought by tracing the origin and development of Islam's political language from the time of Muhammad to the transformation of the Islamic religio-political discourse in modern times (including the etymology of key political concepts in Islam and words such as *jihad*, *aya tollah*, *imam*, *jahiliyya*, *shaykh*, and *fatwa*; see Esposito, 2003; Lewis, 1991). Also, many have chosen to explore the political manifestations of Islam through Islamic organizations pursuing political goals, such as Hamas,

Hizbullah, al Qaeda, Tablighi Jamaat, and Hizb-ut-Tahrir (Ayoob, 2007). Others have focused on the development of a particular stream of Islamic political thought (i.e., political Islam; see Fuller, 2003; Roy, 1998; Tibi, 1998). Finally, some have compared Western and Islamic political thought, noting that key events and thought patterns that have shaped the former, including Europe's "Renaissance" and the gradual secularization of political thought, have not taken place in Islamic political history (Black, 2008).

Conclusion

From its very inception with Muhammad and his life as a religious and political leader of the Muslim umma to modern day Islamic movements, Islamic political thought has undergone tremendous change and development. What started with and from a single individual transitioned to a myriad of thinkers, philosophers, movements, and schools of thought, each of which interpreted the relation between Islam and politics in a unique and often contradictory or controversial way.

This chapter has outlined the many expressions that Islamic political thought has found from its very origins to modern day fundamentalist and (more) liberal Islamic movements. In order to present a more complete picture, key historical developments (such as the religio-political schism within Islam), the contribution of several classical Islamic philosophers, as well as modern works on Islamic political thought have been mentioned. The richness of background material makes it difficult to present a unified and brief summary of the entire evolution of Islamic political thought through the centuries. It suffices to say, therefore, that Islamic political thought has come a long way, and one can only wonder what the next stages of its development may be.

Notes

1. The following works present an in depth study of the pre modern or "classical" period of Islamic political thought: Aslan (2006); Black (2001); Lapidus (1996); and Butterworth (1992).
2. Al Farabi's alternative names and translations from Arabic are *Alfarabi*, *Farabi*, and *Abunaser*.
3. For more in depth study of medieval Islamic political thought, see Crone (2004, 2005) and Lerner and Mahdi (1963).

References and Further Readings

- Ahmad, J. (1994). Ibn Rushd. *Monthly Renaissance*, 4(9). Retrieved May 18, 2010, from <http://www.monthlyrenaissance.com/issue/content.aspx?id=958>
- Al Razeq, A. A. (1972). *Al Islam wa 'usul al Hukum*. [Islam and the principles of government]. Beirut, Lebanon: Al Mu'assassah al Arabiyah Li al Dirasat wa al Nashr.

- Arnold, T. (1965). *The caliphate*. London: Routledge & Kegan Paul.
- Aslan, R. (2006). *No God but God: The origins, evolution, and future of Islam*. New York: Random House.
- Ayoob, M. (2007). *The many faces of political Islam: Religion and politics in the Muslim world*. Ann Arbor: University of Michigan Press.
- Black, A. (2001). *The history of Islamic political thought: From the prophet to the present*. New York: Routledge.
- Black, A. (2008). *Comparing Western and Islamic political thought*. New York: Oxford University Press.
- Butterworth, C. E. (1972). Averroes: Politics and opinion. *American Political Science Review*, 66(3), 894-901.
- Butterworth, C. E. (Ed.). (1992). *The political aspects of Islamic philosophy*. Cambridge, MA: Harvard University Press.
- Butterworth, C. E. (2005). Ethical and political philosophy. In P. Adamson & R. Taylor (Eds.), *The Cambridge companion to Arabic philosophy* (pp. 266-286). Cambridge, UK: Cambridge University Press.
- Corbin, H. (1993). *History of Islamic philosophy*. London: Keagan Paul International.
- Crone, P. (2004). *God's rule: Government and Islam, six centuries of medieval Islamic political thought*. New York: Columbia University Press.
- Crone, P. (2005). *Medieval Islamic political thought*. Edinburgh, UK: Edinburgh University Press.
- Diamond, L., Plattner, M. F., & Brumberg, D. (Eds.). (2003). *Islam and democracy in the Middle East (A Journal of Democracy book)*. Baltimore: John Hopkins University Press.
- Dieterici, F. (Ed.). (1895). *Al Farabi on the perfect state* (R. Walzer, Trans.). Oxford: Clarendon Press.
- El Fadl, K. A., Cohen, J., & Chasman, D. (Eds.). (2004). *Islam and the challenge of democracy: A "Boston Review" book*. Princeton, NJ: Princeton University Press.
- Enayat, H. (1982). *Modern Islamic political thought*. Austin: University of Texas.
- Esposito, J. (2003). *Unholy war: Terror in the name of Islam*. Oxford, UK: Oxford University Press.
- Esposito, J., & Voll, J. (1996). *Islam and democracy*. Oxford, UK: Oxford University Press.
- Fuller, G. E. (2003). *The future of political Islam*. New York: Palgrave MacMillan.
- Hairi, A. (1977). *Shi'ism and constitutionalism in Iran*. Leiden, the Netherlands: E. J. Brill.
- Halliday, F. (1995). The politics of Islam: A second look. *British Journal of Political Science*, 25(3), 399-417.
- Hassan, H. I. (1967). *Islam: A religious, political, social, and economic history*. Baghdad, Iraq: Times.
- Ibish, Y. (1966). *Nusus al fikr al siyasi al Islami* [Text of Islamic political theory]. Beirut, Lebanon: Dar al Tali'a.
- Ismael, J. S., & Ismael, T. Y. (1980). Social change in Islamic society: The political thought of Ayatollah Khomeini. *Social Problems*, 27(5), 601-619.
- Lapidus, I. M. (1996). State and religion in Islamic societies. *Past and Present* (151), 32-37.
- Lapidus, I. M. (2002). *A history of Islamic societies*. Cambridge, UK: Cambridge University Press.
- Leaman, O. (1980). Ibn Bajja on society and philosophy [Ibn Bajja's political philosophy]. *Der Islam*, 57(1), 109-119.
- Lerner, R., & Mahdi, M. (Eds.). (1963). *Medieval political philosophy: A sourcebook*. Ithaca, NY: Cornell University Press.

- Lewis, B. (1991). *The political language of Islam*. Chicago: University of Chicago Press.
- Mahdi, M. (2001). *Alfarabi and the foundation of Islamic political philosophy*. Chicago: University of Chicago Press.
- Mahdi, M., Gutas, D., Abed, S. B., Marmura, M. E., Rahman, F., Saliba, G., et al. (2008). Avicenna. *Encyclopædia Iranica*. Retrieved May 18, 2010, from <http://www.iranica.com/articles/avicenna/index>
- McGinnis, J. (2007). *Classical Arabic philosophy: An anthology of sources*. Indianapolis, IN: Hackett.
- Parens, J. (2006). *An Islamic philosophy of virtuous religions: Introducing Alfarabi*. New York: State University of New York Press.
- Reisman, D. (2005). Al Farabi and the philosophical curriculum. In P. Adamson & R. Taylor (Eds.), *The Cambridge companion to Arabic philosophy* (pp. 52–71). Cambridge, UK: Cambridge University Press.
- Roy, O. (1998). *The failure of political Islam*. Cambridge, MA: Harvard University Press.
- Sherwani, H. K. (1977). *Studies in Muslim political thought and administration*. Philadelphia: Porcupine.
- Sivan, E. (1985). *Radical Islam: Medieval theology and modern politics*. New Haven, CT: Yale University Press.
- Subhi, A. M. (1969). *Nazariyat al Imamah lada al Shi'ah al Ithn'ay 'Ashriyah* [Theory of Imamah according to the Ithna a' Shairya Shi'ism]. Cairo, Egypt: Dar al Ma'aref.
- Tamimi, A. (1997). *Democracy in Islamic political thought*. Retrieved May 18, 2010, from <http://www.iol.ie/~afifi/Articles/democracy.htm>
- Tibi, B. (1998). *The challenge of fundamentalism: Political Islam and the new world disorder*. Berkeley: University of California Press.
- Walzer, R. (1963). Aspects of Islamic political thought: Al Farabi and Ibn Xaldun. *Oriens*, 16, 40–60.

68

CHRISTIAN POLITICAL THOUGHT

MICHELLE D. DEARDORFF

Jackson State University

The classical paradigm of political thought—consisting of Greek and Roman, as well as early Christian, philosophers—is distinct from later philosophical eras because of its communitarian perspective of the state and transcendent source of ethics and morality. Writers in this paradigm argue that the state and political society are necessary for the full development of the individual. For instance, Plato perceives that the state can help men achieve the virtuous life. As with Aristotle, this means that justice and virtue exist only when individuals are fulfilling the societal role for which they are best suited; for Plato in the *Republic*, this occurs when the state assists in such placement. Aristotle also asserts that the state aids in this development through the enforcement of laws. By being forced to behave legally, people become habitually virtuous. Many of these beliefs and values are sustained throughout the Christian phase of the classical era; for such key Catholic writers as Augustine and Thomas Aquinas, the state acts in conjunction with the church for the purpose of sanctifying a sinful and fallen humanity. The state forces the Christian to curb an inherent sinful nature and rest content until the Kingdom of God is fulfilled, even if this control requires the coercion of the heretical into orthodox belief. For classical thinkers, the individual can be fulfilled only within the context of a community.

A second characteristic of the classical paradigm is a shared belief in the transcendency of morals, or a sense of

natural law, overriding the authority of positive law (statutes and policies passed by the government) and the claims of the state. Although there is a great difference in belief regarding from whence such values derive, for these philosophers values emerge from an exterior force. For Plato, the source of ethics is to be found in the Forms, and the Forms are accessed through philosophy and the philosopher's ability to reason, subsequently leading society to understand and pursue these absolutes. Aristotle also views the Forms as the ideal and essence of morals, values, and ethics; he, however, does not believe that they are obtained through philosophy—practical wisdom is the best means of implementing them. Both Augustine and Aquinas perceive that transcendent ethics are based in God as revealed through the Holy Scriptures; however, they too differ as to how humans may avail themselves of these truths. Augustine emphasizes the miracle of revelation, whereas Aquinas places a central focus on the human capacity to reason.

During the medieval era, a transition occurs both in the role of the state and in individuals' relation to it, as well as in perceptions regarding the accessibility of transcendent morals. One response to this paradigm was unabashed secularism, such as found in some of the work of Machiavelli. In the writings of both John of Salisbury and Marsilio de Padua, the role of the state transforms from one of sanctifying and making people more virtuous to the simpler role of maintaining societal order. A second characteristic of

this transition is found in the debate over the source of ethics and values. Whereas for the more traditional writers of the time this source remains firmly centered in God and the interpretation of these values is secure in the hand of the Church, the little individualism discoverable in Augustine and Aquinas regarding freedom of conscience vanishes. The authoritarian state of the medieval era is further strengthened by this power of interpretation and its legitimization by God. Machiavelli, in *The Prince*, also is a part of this strong authoritarian tradition: Power still resides in the ruler to maintain order in society, and the ruler continues as the purveyor of law and ethics.

The source for the demise for the medieval paradigm was the same impetus that led to the creation of the views of the Reformation—both serving as transitions to the liberal political paradigm. A primary inspiration for this vast change was the writings of Martin Luther; although Martin Luther did not directly apply his notions of individualism to the state—he did not believe Christians needed a state but merely obeyed the secular state for the sake of unbelievers in the community—his work is applied directly to politics by another reformer, John Calvin. Luther challenges the authoritarian medieval paradigm through his argument for the priesthood of believers. Individuals, he contends, do not need the intervention of the Church in the relationship between individuals and God. People can govern themselves in their spiritual relationship without requiring an intercessor. Calvin applies this belief in his *Institutes of the Christian Religion* in an effort to return to some of the elements of the classical paradigm. The state and its laws still provide for the virtue and needs of humans for sanctification; however, instead of the philosophers, Church, or society interpreting ethics and morals, individuals can interpret these elements for themselves. These values are then made manifest through the congregational rule of the Church, and the Church implements and enforces these rules through the state.

This Reformation paradigm differed from the medieval view primarily because of its emphasis on the importance and role of the individual in society; although stability and order are still important, these roles do not predominate over other functions of the state. The latter half of the Reformation paradigm is the transition from the Christian theocracy of Calvin and others to the beginning of the secular liberal state. The writings of both Jean Bodin and Richard Hooker illustrate this transition. The source of morals and ethics still derives from the transcendent sources of the Christian faith; however, society is not assumed to be inherently a Christian one. A primary reason for this change in perspective is found in the fact that the Roman church no longer had religious supremacy. The less centrally controlled Protestant church challenged the perspectives of the Catholic state and frequently struggled for control of the throne, as in the kingdoms of England and Scotland. With the emergence of the Calvinist Huguenots in France and the Anabaptists, who explicitly avoided

engagement in politics, tenuous religious equilibrium disintegrated, and the bloody religious wars commenced. The necessity of religious toleration became a key concern of philosophers; with the new world community developing through colonization, a new philosophy of science and new religious and political organizations emerged.

The resulting liberal paradigm emphasizes the prioritization of individual rights over the classical paradigm's communitarian needs. Thomas Hobbes, John Locke, Jean-Jacques Rousseau, and John Stuart Mill all fit into this paradigm because they concur that the role of the state is to protect individual rights and liberties and they believe that sovereignty is inherently located in the free, male, property-holding citizen and permanently ceded or temporarily loaned to the state. God, in this worldview, created government to ensure natural rights of the individual, but does not place sovereignty directly in the hands of the state. The contributions of Christian thought did not cease with the demise of the medieval and Reformation eras; it is clear that elements of the Christian perspective on political philosophy have influenced political thinking, particularly with the framing of the 20th-century expressions of civil disobedience and in the anticolonialist challenges of liberation theology.

Christianity's Emergence in the Political Thought of the Classical World

The famed historian of political thought George Sabine (1961) has noted the following:

[The] rise of the Christian church as a distinct institution entitled to govern the spiritual concerns of mankind in independence of the state may not unreasonably be described as the most revolutionary event in the history of western Europe, in respect to politics and to political philosophy. (p. 180)

While Christianity and classical Roman thought share some basic similarities in their assertions of natural law, human equality, and the necessity for justice, key differences ensured that they would eventually conflict.

First, Christianity makes a claim of egalitarianism that is broader than the Roman assertion of essential human equality. According to the Apostle Paul, in Galatians 3:28, for Christians there could be no distinctions based on ethnicity, the lack of a Jewish heritage, a believer's gender, or whether one was enslaved or free. Second, according to the Christian Messiah, Jesus Christ, the Christian kingdom is not a physical kingdom but a spiritual one (see, for instance, John 18:36). For the Christian, this results in divided loyalties that are the source of much of the original political thought emerging from Christianity. Paul instructs believing Christians in the Roman Empire to be subject to their government, as the King James version translates Romans 13:1–6:

Let every soul be subject unto the higher powers. For there is no power but of God: the powers that be are ordained of God. Whosoever therefore resisteth the power, resisteth the ordinance of God: and they that resist shall receive to themselves damnation. For rulers are not a terror to good works, but to the evil. Wilt thou then not be afraid of the power? Do that which is good, and thou shalt have praise of the same: For he is the minister of God to thee for good. But if thou do that which is evil, be afraid; for he beareth not the sword in vain: for he is the minister of God, a revenger to execute wrath upon him that doeth evil. Wherefore ye must needs be subject, not only for wrath, but also for conscience sake. For this cause pay ye tribute also: for they are God's ministers, attending continually upon this very thing.

Christ, however, encourages his followers to give to the political state what it requires while simultaneously remaining loyal to the demands of God. In Matthew 22:17–21, debating with some of the Jewish leadership, Christ responds to the query, “Tell us therefore, What thinkest thou? Is it lawful to give tribute unto Caesar, or not?”

But Jesus perceived their wickedness, and said, “Why tempt ye me, ye hypocrites? Shew me the tribute money.” And they brought unto him a penny. And he saith unto them, “Whose is this image and superscription?” They say unto him, “Caesar’s.” Then saith he unto them, “Render therefore unto Caesar the things which are Caesar’s; and unto God the things that are God’s.”

Unlike Roman philosophy, in which the gods expected citizens to owe loyalty to the Emperor, for the Christian, it was only to the office of the ruler that citizens owed allegiance, not to the specific individual. Although the question of obligation to an unjust ruler is not new (consider, for instance, the Greek playwright Sophocles and *Antigone*), this tension is embedded within Christianity. For the Christian, unlike the classical pagan, this religion has a higher calling on the individual than merely living a virtuous life as a citizen of the state. In fact, Christianity places a calling on an individual’s life more powerful than merely the duty of civic obedience, demanding commitments from the individual that no earthly sovereign can eradicate. This tension created an inherent conflict with the Roman Empire, resulting in persecution of these dissidents. The attempted destruction of Christianity in the latter part of the 3rd century was justified, not by Christianity’s religious competition with paganism, but by its alleged attempts “to build up a state within a state, its boring from within every social class, and its gradual absorption of the Roman empire by infiltration and ideological appeals without overt acts of force” (Ebenstein & Ebenstein, 2000, p. 182).

Christian persecution ended after the rule of the Emperor Diocletian in 303–305 CE; by 312 CE, the Emperor Constantine personally converted to Christianity. One year later, he endorsed the Edict of Milan, ending the

persecution of Christians, guaranteeing the freedom to profess the faith without any fear of state involvement, and formally recognizing the Christian church. This edict resulted in a new political climate in which Roman rulers sought Christian support for their policies. As a relatively newly institutionalized religion, Christianity experienced much internal tension over questions of doctrine and creed. In 325 CE, the Council of Nicaea settled many of these primary conflicts within Christianity. This determination identified which beliefs were to thereafter be classified as heresies and what schisms would emerge in the Church as a result of this delineation of orthodoxy. Although the Council ensured that Christianity survived as a coherent set of beliefs and doctrines, it also resulted in the persecution of many groups that demurred to the orthodox views of Christianity; most particularly, these controversial issues focused on *Christology*, or the nature of the Christ. By 393 CE, the Emperor Theodosius had declared Christianity to be the official religion of the Roman Empire, resulting in the birth of the Holy Roman Empire; by 410 CE, however, the Visigoths, under the leadership of King Alaric, invaded for the third time, finally looting the riches of Rome.

Within the growing Christian church were two challenges that required addressing. For many citizens of the Roman Empire, especially those who worshipped the traditional gods, the loss of Roman control and authority was traceable to the ascension of Christianity with the Empire. In addition, once the Council of Nicaea had explicitly identified heresies and schisms, intense tensions divided those who followed the now orthodox Christianity and those who retained beliefs now deemed heterodox.

Augustine (354–430)

Augustine, one of many African bishops within the Roman Empire, addressed these internal stressors on the Christian Church. His most famous book, *Civitas Deis*, or *City of God*, was largely designed to demonstrate that Christianity was not responsible for the demise and invasion of Rome. Instead, he considers how believers can live with the demands of the state while simultaneously pursuing the requirements of an obedient Christian life. Augustine’s philosophy bridged classical Greek and Roman thought with Christianity, particularly integrating Platonic thought and values with a Christian worldview. More specifically, the Platonic understanding of justice was immersed in the ideals and values of Christianity. Augustine believed that Platonic thought was very similar to the Christian construction of the world although there were clear conflicts between the two. In Augustine’s work, classical thought was transformed.

Humanity’s rejection of God in favor of self is the key turning point for Augustine. The badge of sin carried by every newborn child—natural depravity—means that the world is a sinful place, existing outside of God’s original

plan; consequently, the ideals of justice cannot be realized on earth. Each person must make a personal decision in response to the reality of original sin. Based on this choice, Augustine divides the citizenry into two groups—those who choose to live in the City of God and those who select the City of Man. Although they are not physical cities and although no one can ever know with certainty where any individual truly resides, these demarcations indicate that people in a sinful world will choose to place their allegiance and priorities either in loving and serving God or in loving and serving themselves. Physical membership in the Church, according to Augustine, is not an accurate indicator of true citizenship. Augustine argues that while ideally an earthly ruler will be a denizen of the City of God, realistically most monarchs will be citizens of the City of Man, even if they claim otherwise. The only way to potentially uncover where a person has chosen to place his or her values is by watching how the person lives, and even then an observer might be inaccurate. This understanding that true Christianity is purely internal would later resonate throughout the Reformation.

Arlene Saxonhouse (1985) notes that because in the City of Man the body, not the spirit, is dominant, women, as others who are politically oppressed, are inferior. In the City of God, woman can be equal—her soul is equal to man's because both have a direct relationship with God. The City of God removes the need for women to perform within the context of the family, allowing for a more true equality. But in Augustinian thought, this equality was in existence only in the City of God, not in the City of Man. Similarly, slavery is a consequence of the sinful world, not a natural phenomenon constructed by God. For the vast majority of classical and medieval philosophers, women, slaves, foreigners, children, and servants are dependents entrusted to the “citizen” to be protected and used. Therefore, with a few exceptions, as in Augustine's City of God, they were not considered to be theologically or politically relevant to these philosophical paradigms.

Because of humanity's fall from God's grace and humans' natural depravity, in which humans reject God's will and embrace their own, Augustinian thought asserts that individuals are wholly incompetent at governing themselves. The only hope of human freedom is found through service to God, manifested imperfectly on Earth through the Church. Consequently, believers require guidance to help them remain obedient to God's commandments, which are partially communicated through the Church and the Scriptures. According to Augustine, God uses the state to compel obedience. The form of governance is irrelevant; obedience is due to *any* earthly government because God makes human beings dependent on both the Church and the state. The function of the state is to provide social peace, albeit one that is imperfect and temporary, because the service and obedience required by God are possible only in an ordered and peaceful society. The state protects humans from themselves and assists in

their sanctification (the process of becoming more like Christ) and moral maturity. Justice, however, is impossible to achieve on Earth; peace and stability are the best for which we can hope. For Augustine, compelled obedience also helps the heretical become more orthodox. Because God is the source of all truth, reality, and morality, God's will is best relayed to the people through the Church and, when the state is obedient to God's commands, communicated through the state. When the Church is obedient, compelling citizens to be obedient to the Church and state, citizens must obey both. But even if the state is disobedient to the rule of the Church and to the commandments of God, believers cannot deny the authority of the state. If Christians are commanded to go against the word of God, however, they should be willing to die as martyrs to the faith rather than be subversive to the state. For Augustine, good, obedient citizens (orthodox Christians) have nothing to fear from the state and therefore no reason to rebel. The state is used by God to aid in the growth and sanctification of the good, punish the evil into reformation or destruction, and move the heretical into orthodoxy. This is one of the earliest statements, endorsed by the Roman Church, as to the proper relationship between Church and state.

John of Salisbury (1115/1120–1180)

Near the end of the 5th century, Pope Gelasius I defined the frequently contested relationship between the ecclesiastical power of the Church and the secular power of the state in a manner that would be later described as the *two swords* formulation. Although Christ is both the prince and the pope, according to Gelasius I, Christ divided these offices to protect humanity from itself, giving to the Church the responsibility of the spiritual welfare of the people and to the state the administration of secular politics. Both rulers derive their authority directly from God, yet according to this model, each office is independent and sovereign in its own realm. By the 12th century, this cooperative model would be strongly challenged by the papal authorities in the Church, who argued that God had given all earthly authority to the Church, which then delegates political power to the state.

John of Salisbury provides a typical medieval articulation of this papal position and is one of the earliest attempts at a coherent political theory in the Middle Ages prior to the Western rediscovery of Aristotle's scientific writings in the 13th century. In his work *Policratus* (*The Statesman's Book*), John of Salisbury constructs a sophisticated comparison of the physical body to the republic, in which the different elements of society are identified as equivalent to the body politic. For instance, the military serves as the hands of the body, while the bureaucratic agencies act as the internal organs. The Church is the soul of the body or of the republic—not separate from it. It is

the Church that *provides* the sword to the prince, with the caveat that the Prince not exploit or destroy the clergy. The purpose of the state is to protect the Church and clergy from injury both from itself and from the state and to maintain order within the people. God granted power to the Church, which then delegates physical authority to the state, which must remain responsible to the Church. This does not mean that the papal authority has a veto over the choices of the prince, nor does it require that the Church control the state, but it does require that a governmental statute or ruling be nullified if it conflicts with the teachings of the Roman Church.

A second theme in the *Policratus* is John of Salisbury's recognition of the potential for abuse by wielders of both swords and the evidence of said abuse in the contemporary Roman Church. Because he had much practical experience in politics, including serving as secretary of two Archbishops of Canterbury and possibly witnessing the murder of Archbishop Thomas Becket by King Henry II's assassins, John of Salisbury understood the consequences of this abuse. In the *Policratus*, he accuses the Church of having greedy priests who exercise duplicity in their agendas, manifest a lust for power, and demonstrate a lack of compassion for those who suffer. Because of the eternal consequences of such abuse by the ecclesiastical authority, he argues that a tyrant in the Church is worse than a secular tyrant of the state.

This work is probably best known for its argument, unique in the Middle Ages and prescient of John Locke's *right to rebellion*, that there can be legitimate grounds for citizens to destroy a tyrant. For John of Salisbury, there are laws that even kings can become outlaws for breaking; tyrannicide is acceptable when the ruler violates certain laws, particularly those regarding the authority of the Church. There is a mutual obligation to law binding the ruler and the ruled, so that the distinction between legitimate ruler and tyrant is essential. Although tyrannicide is permissible, it is essential that the citizen pursuing the deed distinguish between the appropriate consequences for crimes and vices, that tyrannicide not be committed by someone who has made a sacred oath to uphold the ruler, that the citizen respect the biblical injunction against poison, and that the citizen realize that tyrants can be used by God to punish those who are evil and to discipline the good. John of Salisbury was one of the few Roman Church authors who legitimated the disposal of God's ordained, even though he limits this remedy to specific circumstances.

Thomas Aquinas (1225–1274)

While Augustine is understood as integrating classical Greek and Roman thought into Christianity and in particular connecting Platonic thoughts and values with a Christian worldview, Thomas Aquinas is recognized for

applying Aristotelian logic and systemic thinking to Christian doctrine. In Christian political thought, only Aquinas parallels the impact of Augustine, both integrating classical thought into Christian theology. As famously stated by William and Alan Ebenstein (2000), "To be born, the Church needed Plato. To last, it needed Aristotle" (p. 222). Initially understood through early Arabic and Jewish commentaries, Aristotle's *Politics* and *Ethics* were translated from Latin into Greek during Aquinas's lifetime. Aristotle's scientific writings had a tremendous impact on medieval political thought. Many Christians, as Jewish thinkers had done centuries earlier, attempted to fit Aristotelian thought into a holistic synthesis of scientific and theological understanding. In this process of synthesis, however, Aristotle is reinterpreted and transformed. Writing in the midst of the rediscovery of Aristotle and the debates over the role of medieval law, Aquinas mediates between the Aristotelian presumption that human reason can help obtain justice and the Church's assertion that the basis of right is custom and tradition. Aquinas tries to integrate both custom and reason, legitimizing both king and pope. The king can rule, but only where law is supreme and the King pursues justice. The Aristotelian function argument remains in Thomist thought (as the philosophy of Thomas Aquinas is generally referenced)—happiness is the end (meaning purpose) of the state, if all fulfill their societal roles and positions, when this is achieved and the king is subordinate to both the Church and God, Aquinas believes justice may exist. While Aquinas generally holds to Gelasian assumptions regarding the dual authorities of Church and state under the two-sword theory, he asserts that under specific conditions, the Church can remove a prince and release citizens from their political obligations to the ruler. The primary political works in which this analysis is developed are *On Princely Government*, *Of the Governance of Princes*, and *Of Rulership* (also known as *On Kingship*, which has been historically questioned as to its authenticity).

The *Summa Theologica* is the best example of Aquinas's attempt to fully integrate the competing claims of faith—demonstrated by the Roman Church's theological doctrine—and reason—best articulated by Aristotelian thought. For Aquinas, faith and reason both derive from God, and thus they can never truly be in conflict; faith, however, is a direct communication from God and is therefore closer to truth. Each section of this massive work is presented in a parallel structure beginning with a question under contemporary theological discussion, followed by "objections" reflecting the relevant erroneous answers to the question. Aquinas then provides the doctrinally correct answer to the question, generally supported by quotations from such authoritative sources as the Bible. Additional supportive evidence is included, and each numbered section concludes with a specific response to each of the original objections demonstrating its fallacies. The *Summa Theologica* provides definitive doctrinal

answers to the key questions of the day; although alternative positions are evaluated, reason, supported by faith, reveals the truth. The Roman Church later adopted this massive tome as the authoritative statement of Church doctrine, used to teach young ordinands as they entered the priesthood.

Unlike Augustine, in Thomist thought the individual Christian can seek justice in both this world and the next. Obedience to authority is a virtue; a temporal ruler can legitimately expect the people to obey him; and through this obedience, citizens exchange their skills and talents for security and peace. So while an individual requires a state and is subordinate to it, the individual can expect something in return—security, peace, and, ideally, justice. Contrary to Augustinian thought, the function of government is to create a more just, secure society in which people can find happiness; justice is defined as individuals' receiving their due by virtue of their contributions to society. According to Saxonhouse (1985), Aquinas believes that hierarchy is a reflection of the eternal order among humans. Kingship is always the best part of government; thus the subordination of the female is part of the order of nature as well. While for Augustine this oppression, like slavery, demonstrates the inadequacies of a corrupt world, for Aquinas it is simply part of the natural world ordered by God.

To prevent tyranny, Aquinas believes, the ruler must govern within the constraints of the law. Consequently, a role of the Church is to remind the ruler of his limitations through the threat of excommunication—for, consistent with the Roman philosopher Cicero, without law, there is no justice. The ruler has the right to expect people to obey him because within obedience to the law, they can expect an imperfect justice. Custom and tradition, however, are subject to criticism through the God-granted facility of reason; if the customs and traditions are unreasonable, they can be questioned and rejected. This freedom to question and challenge marks the early origins of a liberal view of individualism. Although Aquinas recognizes the dangers of tyranny, Thomist thought reflects the assumption that revolution in response to tyranny can result in worse abuses. For this reason, as well as the ruler's direct appointment by God, Aquinas rejects John of Salisbury's endorsement of tyrannicide.

For Aquinas, law has multiple manifestations. Because laws emanate from God, they are rooted in the universal and are applicable to all cultures, across time and circumstances. In Thomist thought, there are four types of law: the eternal law of God revealed in the universe, the divine law of God communicated through the Scriptures and the edicts of the Church, the natural law of God understood through the experiences and realities of humanity, and the human law through which eternal values and expectations are translated into legislation. For Aquinas, the key elements of natural law are (a) natural inclinations such as self-preservation, (b) engrained instincts such as procreation

and the education of children, and (c) the internal propulsion of human beings to reason, toward knowing God and His truth, as well as to life in community. This formulation of natural law remained constant throughout the European Enlightenment.

Dante Alighieri (1265–1321)

Similar to Machiavelli, albeit born 200 years earlier, Dante Alighieri lived in Italy as the numerous Italian states battled for dominance. As feudal society transitioned to a world of independent cities, the Church began losing control over local governance; it is not surprising that philosophers and scholars would seek better, more constant forms of governing. Although his *Divine Comedy* is better known, Dante's *De Monarchia* (circa 1310) is understood to be one of the most important challenges to centralized papal powers in the Middle Ages. His work is interpreted as a unique combination of both Augustinian and Thomist/Aristotelian thought.

Dante makes three primary arguments in *De Monarchia*, all regarding the proper governance of society and the ensured thriving of humanity through sustained political associations. In the first portion of his work, Dante argues that in order to have the uninterrupted peace necessary for humanity to develop to its full potential, a universal monarchy is required. Unity is essential to guarantee that states can resolve their disputes without resorting to war. A monarch can best secure the freedom necessary for individual and communal development. For Dante, however, this monarch would resolve only matters that require a common rule; most issues would be reserved for the sovereignty of the local state or community traditions. There is some debate over whether Dante intended to advocate for a worldwide monarch (Ebenstein & Ebenstein, 2000) or simply to unify Italy (d'Entreves, 1952).

In his second argument, Dante recommends that the nature of this universal government should be Roman because the Roman Empire acquired its domination of the world through natural right and its divine appointment by God, ruled based on law, achieved the common good for all, and ensured peace with liberty. It is this combination of peace and liberty that Dante believes will ensure human society can fulfill its potential. His final argument addresses the appropriate relationship of the Church with the state. For Dante, unlike many of his compatriots, the authority of the emperor is delegated directly from God and is independent of the intermediary of the pope. As a human is both an earthly and a spiritual being, possessing both sets of attributes, governing bodies must have both essences. To have a blessed earthly life, reason and philosophy as articulated through human law can be protected by an ordained emperor; to achieve the heavenly paradise, people must move beyond human reason to faith guided by the Church. By necessity, individuals need two

guides—the pope to lead the citizenry to eternal life and the emperor to guarantee earthly happiness.

Marsilio de Padua (1275/1280–1342)

In 1302, Pope Boniface VIII (1294–1303) issued the papal bull *Unam Sanctam*, which stated that only the Christian Church provided the means through which salvation and the forgiveness of sins occur. This Church had two swords—one spiritual and one secular—but “both swords are in the power of the church, the one by the hand of the priest, the other by the hand of kings and knights, but at the will and sufferance of the priest.” *Unam Sanctam* noted that the highest temporal authority could be held accountable by the spiritual power of the Church but that only God could judge the highest spiritual authority—the pope. A culmination of many battles between secular rulers and the Church, this papal bull resulted in outright and successful rebellions by such monarchs as England’s King Edward I and France’s King Philip IV.

Marsilio de Padua’s work, *Defensor Pacis* (*The Defender of the Peace*), in 1324, is significant because it makes a “positivistic separation of laws and morals, [establishes] civil power on nontranscendent grounds, and [deposits] political authority in the people as a whole” (McDonald, 1968, p. 176). This work marked the beginning of the secularization of the state, in which citizens—not God—are the source of governing legitimacy; this move toward the modern conception of the secular state is often attributed to Machiavelli but is traceable to Marsilio. Explicitly building on an Aristotelian comprehension of the origin and role of the state, Marsilio is led to a conclusion different from Aquinas’s regarding the authority of the Church, although all three philosophers conclude that the role of the state is to provide the good life. This “good life” has two components for Marsilio: the use of philosophy via reason to secure the good life temporally and to use revelation via faith to have the good life in the eternal realm. Consequently, like Dante, there is a need for both civil and religious government. The citizenry grants authority for this civil government, and many commentators (but not all; see Strauss, 1987, p. 284) perceive this idea as an explicit statement of popular sovereignty, albeit with the exclusion of women, children, foreigners, and slaves. The common will of the people is the source of political authority, and this will is known as the *Legislator*. The agent of the Legislator—the ruler—is the executive of the government. In Marsilian thought, this ruler is an elected monarch, although not inherently an individual. Although there may be divine law, it is human or positive law that possesses legitimacy. Divine and human law are distinguished from each other by the nature of their penalties when they are trespassed; if penalties are eternal, such laws cannot be enforced on Earth, and if laws are temporal, all are accountable to

them, including king and priest. If the king violates the laws, the Legislator (corporate citizenry) is able to hold the King accountable, as with any citizen. In making this distinction between human and divine law, Marsilio de Padua attacks papal power and argues that the Church must be subject to secular judges. He removes all coercive power from the hands of the Church, not, as some assert, to allow for religious freedom, but to clearly distinguish enforceable positive law from the divine law realized only by God.

Marsilian thought is not a devaluation of religion or Christianity. Marsilio argues that the activity of the Christian priest is the most noble act of any believer, but Marsilio also articulates concerns regarding the corrupting influence of the power of coercion on the Church. By destroying ecclesiastical hierarchy, finding no authority for this power in the Scriptures, Marsilio places the individual priest and the corporate body of the Church under the authority of the state, just as every other individual and corporation is under its authority. This destruction of papal imperialism and the challenge to Church corruption anticipate the concerns of the Reformation.

Martin Luther (1483–1546)

The conciliar movement was an attempt by the Roman Church to address its widely perceived corruption by giving decision power previously assigned solely to the pope to councils that had authority to reform Church structure. Two councils were convened—Constance during the period 1414 to 1418 and Basel from 1431 to 1439—but neither was effective in advancing systemic change. When the ecclesiastical structure was unable to reform itself, a revolt from the membership of the Roman Church was inevitable. While Martin Luther was neither the first nor the last to advocate theological and political reform, he was the most influential in both instigating and fulfilling the Protestant Reformation. His famed declaration in the Ninety-Five Theses (1517) was his initial attack on papal indulgences, which he and others believed had corrupted the Church both theologically and politically. This system of indulgences instilled by elements of the Western Church had guaranteed salvation to those who could afford it, while enriching the priesthood and impoverishing many believers. The Protestant Reformation drew heavily on the theological arguments initially made by Augustine in the 5th century, challenging the Roman Church’s more recent reliance on Aristotelian thought.

Luther’s primary political works are *Treatise on Christian Liberty* (1520) and *Secular Authority: To What Extent It Should Be Obeyed* (1523). Luther’s religious and political contributions are parallel: Individuals can understand God’s word directly in an unmediated relationship. Instead of the laity requiring a dedicated priesthood to

intervene with God, Martin Luther argued for a “priesthood of all believers.” Embedded within this schema is a notion of basic equality of all believers as Christians, but as with Augustine, this equality does not translate into a temporal format. Neither the Church nor the state is required to intervene within the relationship between humans and God. As a corollary, because faith is inherently personal and internal, true belief can never be coerced, only right behavior. Despite the claims of the Aristotelians, believers should not seek religious truth through reason—although individuals have the capacity to reason—but through their capacity for belief.

The question of the best form of government is mostly irrelevant to Luther, because God provides government for the guidance of the sinful person. Luther appears to support a monarchy above other forms because he fears any form of democracy would inevitably result in mob rule and control by the wicked. Christians themselves need no laws because they are governed directly by God; however, they obey and support government for the sake of their nonbelieving neighbors. In the works of Martin Luther, government is not religious in nature (unlike the perspective of John Calvin), and Luther does not perceive a Christian state to be feasible. Following Augustinian thought, government is ordained for a sinful world; therefore, a Christian government is impossible because evil always outweighs the good in the temporal sphere. The purpose of the government is to provide order, and therefore earthly justice should not be expected. While God has ordained two kingdoms, one religious and one temporal, they must be independent of one another. Luther’s key concern is that preaching of the Scripture, offering of the sacraments, and interpreting of doctrine are protected from the power of the state.

God is the source of all ethics and morality, and his will is directly revealed to individuals through faith; this is a process entirely separate from governance. While there is no right to rebellion, the believer does have a right of passive resistance. The individual does not have to obey despite conscience or faith’s dictates; there is no personal necessity to follow an evil ruler’s wrong edicts or to embark on an unjust war. But, for Luther, this view does not legitimate any form of organized resistance. It allows only for *passive resistance*, a personal response to evil rule. In fact, as Luther’s thought developed, his attitude toward dissenters and rebels hardened. In 1525, for instance, the German peasants who had been heavily oppressed both politically and economically took Luther’s religious theories quite literally and revolted. Luther immediately supported the princes in brutally crushing rebellion. Although he recognized the unfairness of the policies that had been enforced by those in power and that had informed the revolt, for Luther obedience to rulers is still the duty of believers because, as he argues, the world is a wicked place and deserves such harsh governance.

John Calvin (1509–1564)

John Calvin was a French Protestant who moved to Switzerland, a newly Protestant country, because of religious oppression. There he wrote *Institutes of the Christian Religion* (first edition, 1536) and governed Geneva (1536–1538 and 1541–1564) in an attempt to realize his perfect Christian society. The ideal government for Calvin is a theocracy; government is good, provided by God, and the state should support the Church. Obedience rendered to the state thereby equals obedience rendered to God. As with Luther, Calvin perceives two types of government—the spiritual and the political—that complement and assist each other. His city of Geneva was to install this new world order and provide moral guidance for citizens; consistent violators of this order were to be expelled from the community. Geneva brought discipline to people displaced and excluded in the old system, thus creating a larger, better functioning workforce.

Calvin asserts that the function of government is to make people moral by providing order and justice to the larger civil society, believing that obedience to God’s law leads to justice among his people. Laws should not neglect God, but the state’s primary function is to aid the Church by enforcing laws with the objective of making people virtuous. As with other Christian philosophers, Calvin understands that God’s will is revealed to those who have a direct relationship with him, but the Church enacts God’s will through the enforcement of the state. Individuals need to be protected from corrupt societies that avoid teaching morality or reject orienting their members into right values; such societies need reform.

For Calvin, the believer seeks success, which is defined as material benefits to be saved, not enjoyed. Such monetary savings are understood as a social resource, a foundation of the industrial revolution; this expected austerity is a virtue of self-control and prevents the rule of lust in the life of the individual. Calvin is clear that the essence of the work ethic is found in the individual; like Aristotle, virtue is enforced by the law, and people slowly become virtuous through habituation perpetuated by law. This value was possibly best exemplified in the culture of Geneva and in the Puritan societies in colonial Massachusetts. An immense debate exists in the literature over the role of Calvinism in the creation of modern capitalism (see Green, 1959; Weber, 1930) by its removal of theological barriers to a capitalist system.

More than Martin Luther, Calvin recognizes that some resistance is acceptable in the case of tyranny. God is sovereign and will hear the cries of His people and deliver them by a savior. The purpose of the magistrates, for instance, is to check the power of rulers, and the magistrates should exercise this authority. Lesser governmental officials have a duty to protect the political sphere from a tyrannical leader; their right to resist comes from God because the sovereign power is shared. In a good system of

government, the prevention of tyranny should be automatic because authority is divided and there are automatic checks on the consolidation of power. While Calvin emphasizes obedience and not direct resistance, his followers transformed this reasoning. Huguenot interpretation of a thread of Calvinist thought led to such essays as *A Defense of Liberty Against Tyrants* (1579), published under the name of Stephen Junius Brutus, which justified a contractual understanding of government, popular sovereignty, protection of property, and the right of some resistance against tyrants. Similarly, John Knox rejected Calvin's notion of passive resistance, arguing to the Scottish Protestant church that it is the duty of believers to challenge and resist a king who behaves contrary to God's word and God's glory.

Jean Bodin (1530–1596)

In his *Six Books of the Commonwealth* (1576), Jean Bodin creates a modern notion both of the state and of sovereignty. The family is the basis and the origin of the state—resulting in a strong distinction between public authority of sovereigns and private authority of heads of households. The ruler has been granted absolute and perpetual power under God and thereby has an immense obligation to serve him. Consequently, Bodin believes that divine retribution will fall on evil rulers. Sovereigns, however, do not have to be kings. They may be either individual or collective in their composition. While Bodin prefers a monarchy, he argues that legitimate sovereignty can be manifested in any form of government. The duty of the state has long been to protect property; it is not viable for the modern state to unify public and private happiness because the modern state is big, diverse, pluralistic, and must be ruled by a dominant central power. Bodin provides a mix of the old and the modern, but his work marks the end of the concept of the unified Christian society. While he rejects much of Calvin's and Luther's analysis regarding the interrelationship of Church and state, he advances the supposition that religious belief is a personal and not a public concern by explicitly advocating religious tolerance by the state.

Richard Hooker (1554–1600)

Like Bodin's work, Richard Hooker's *Of the Laws of the Ecclesiastical Polity* is a link between the medieval and modern conceptions of government. Hooker, although a Protestant, still values tradition, authority, good order, and law but also manifests a high degree of tolerance for religious dissent. The law of nature, however, requires that people have some kind of government or governing structure. At root, the government of the Church and the state are one, but they are not controlled by an all-powerful authority. The sovereign, unlike for Bodin, is not the one

whose will becomes law, but instead the sovereign exists to enforce preexisting law. The sovereign is the “King-in-Parliament,” not the king as an isolated, independent ruler. Monarchy is not an absolute form, but if a monarch rules and the society is Christian, then the monarch must be Christian. Ethics and truth still are provided through natural law and God's revelation through Scripture. Individuals, as in the view of most of the medieval philosophers, are still denied the right to resist tyrants because although tyranny is very destructive, anarchy is much worse.

Some Contemporary Manifestations

The contemporary impact of Christian political philosophy has been seen across the political spectrum in Western societies. The vision of the Christian reconstructionist who wishes to return to a literal Old Testament legal system and the perspective of the Black Liberation theologian who views the New Testament priorities of the Sermon on the Mount as providing a systemic definition of legal justice—both derive their impetus from the political philosophies of the past. Contemporarily, there are three political interpretations that have been quite pervasive in Western thought: the vision of the Anabaptists, civil disobedience, and liberation theology.

Anabaptists

An additional response to the new theological and political assertions of the Reformation is found in the development of Anabaptist communities in which the response to a personal God central to the believer's life is a rejection of the corrupting influence of political engagement. Direct descendents from the radical reformers of the Protestant revolution, current denominations that derive their theological stances from Anabaptist premises include the Amish, the Mennonites, the Church of the Brethren, and Hutterites. The term *anabaptist* derives from the Greek word that means *rebaptize*, reflecting the understanding of baptism as a sacrament in which only believers could partake and rejecting the pervasive Catholic and Protestant acceptance of infant baptism. Although there are great theological differences among these communities, they also hold some basic premises in common, most clearly articulated in the Schleithem Articles of 1527. They are generally pacifistic, refusing to bear arms or to serve in the military, and believe in both nonviolence and nonresistance. Anabaptists endorse the strict separation of church and state because they do not believe that the state can supersede the requirements of God's law and the church must be free to worship independent of state regulations. To different degrees, Anabaptist communities withdraw from the larger secular society in order to be more pure in their relationship to God.

Civil Disobedience

While civil disobedience certainly does not have its roots in the Christian Church, many of its practitioners have justified their participation within their Christian faith. Civil disobedience is an attempt to challenge the legitimacy of a country's laws and practices without contesting the legitimacy of the nation. By nonviolently disobeying laws and passively accepting governmental consequences, activists hope to call attention to the injustice of the policies they are challenging. Many practitioners, such as Martin Luther King Jr., based their justification of this practice on the scriptural contention that humans are to obey God's law and disobey human law when it is unjust. In his famed "Letter From a Birmingham Jail," King notes that he agrees with Augustine's claim that an "unjust law is no law at all" and Aquinas's distinction that "an unjust law is a human law that is not rooted in eternal law and natural law."

Liberation Theology

Liberation theology emerged from impoverished colonized communities in Latin America and was built on such writings as Peruvian Roman Catholic theologian Gustavo Gutiérrez's *Theology of Liberation* (1971). It insists on the centrality of the praxis (practice) of Christianity. This perspective recognizes the challenging of oppressive political systems as central to the doctrine of Christianity, and it privileges the experiences and voices of the poor as the distinctive of the faith. Because a sinful world is the root cause of poverty, only through the institutional challenging of political and economic systemic oppression does the Church pursue God's will in confronting and defeating sin. Godly practice requires that social policies grant preferential treatment of the poor. Prior to ascending to the papacy as Pope Benedict XVI, Cardinal Joseph Ratzinger wrote a refutation of the aspects of liberation theology dependent on Marxist interpretation of history and economics and that are supportive of social revolutions. He noted that although critiques by liberation theologians of the history and theology of the Catholic Church were often accurate, solutions solely dependent on Marxist analysis border on the heretical and challenge orthodox thought. Despite these challenges, liberation theology's interpretation of Christianity has been incredibly influential, not only in Latin America but also in Asia and Africa, and particularly outside Latin America within Protestant denominations. In

North America, its offshoots include feminist liberation theology and Black liberation theology.

References and Further Readings

- Bouwsma, W. (1988). *John Calvin: A sixteenth century portrait*. London: Oxford University Press.
- Cochrane, C. N. (1940). *Christianity and classical culture: A study of thought and action from Augustus to Augustine*. New York: Oxford University Press.
- Deane, H. A. (1966). *The political and social ideas of Saint Augustine*. New York: Columbia University Press.
- d'Entreves, A. P. (1939). *The medieval contribution to political thought: Thomas Aquinas, Marsilius of Padua, Richard Hooker*. London: Oxford University Press.
- d'Entreves, A. P. (1952). *Dante as political thinker*. London: Oxford University Press.
- Ebenstein, W., & Ebenstein, A. (2000). *Great political thinkers: Plato to the present* (6th ed.). New York: Thomson Wadsworth.
- Finnis, J. (1998). *Aquinas: Social, legal, and political theory*. Oxford, UK: Oxford University Press.
- Green, R. W. (Ed.). (1959). *Protestantism and capitalism: The Weber thesis and its critics*. Boston: D. C. Heath.
- Gutiérrez, G. (1971). *A theory of liberation: History, politics, salvation*. Maryknoll, NY: Orbis.
- King, M. L., Jr. (1964). Letter from a Birmingham jail. In M. L. King Jr., *Why we can't wait*. New York: New American Library.
- MacCulloch, D. (2005). *The Reformation: A history*. New York: Penguin Books.
- McDonald, L. C. (1968). *Western political theory: Ancient to medieval*. New York: Harcourt Brace Jovanovich.
- O'Donovan, O., & O'Donovan, J. L. (Eds.). (1999). *From Iranaeus to Grotius: A sourcebook in Christian political thought*. Grand Rapids, MI: Eerdmans.
- Sabine, G. H. (1961). *A history of political theory* (3rd ed.). New York: Holt, Rinehart & Winston.
- Saxonhouse, A. W. (1985). *Women in the history of political thought: Ancient Greece to Machiavelli*. New York: Praeger.
- Strauss, L. (1987). Marsilio de Padua. In L. Strauss & J. Cropsey (Eds.), *History of political philosophy* (3rd ed., pp. 276–295). Chicago: University of Chicago Press.
- Thompson, C. (1984). *The political thought of Martin Luther*. Brighton, Sussex, UK: Harvester Press.
- Walzer, M. (1982). *The revolution of the saints: A study in the origins of radical politics*. Cambridge, MA: Harvard University Press.
- Webb, C. J. (1932). *John of Salisbury*. London: Methuen.
- Weber, M. (1930). *The Protestant ethic and the spirit of capitalism*. London: Allen & Unwin.

69

EARLY MODERNS AND CLASSICAL LIBERALS

DAVID SCHULTZ

Hamline University

Modern political thought is punctuated by five major revolutions, each of which helped forge a break from the ancient and medieval worlds and which have shape politics and society into the 21st century (Blumenberg, 1987b; Tannenbaum & Schultz, 2004). These revolutions are the Protestant Reformation of the 16th century, the Copernican scientific revolution of the 16th century, the French Revolution of 1789, the Kantian revolution in philosophy (marked by the publication of Immanuel Kant's *Critique of Pure Reason* in the 1780s), and perhaps most important for politics, the English Glorious Revolution of 1688 and the publication of John Locke's *Two Treatises of Government* in 1689–1690 (Ashcraft, 1986). Collectively, these five revolutions ushered in a world politically far different from the one that preceded it. The politics of the ancient and medieval eras was marked by an emphasis on absolute governments that expressed little concern for individual rights and vested unchecked authority in monarchs who often invoked claims of superior knowledge, religion, or appeals to familial metaphors to justify their authority (Strauss, 1953). But the world of modernity was a secular one that constructed limited governments based on individual consent and respect for rights as the cornerstone of political power (Baumer, 1977; MacIntyre, 1984; Riley, 1982; Strauss, 1953).

To tell the story of the emergence of modernity—of the Enlightenment and early moderns and classical liberals—is

the task of this chapter. It first contrasts the world of the ancient and medieval eras to that of what has been called modernity. The chapter then describes the five major revolutions that shaped the rise of modernity and the emergence of a new liberal world order that broke decisively with the one that it replaced.

Challenging the Medieval World: The Protestant Reformation

The first major challenge to the medieval world was the Protestant Reformation in the 16th century, which found its origins in the writings and political movements surrounding German theologian Martin Luther (1483–1546) and the Swiss John Calvin (1509–1564). Both spearheaded what could be described as conservative reactions to philosophical reforms and excesses that had emerged in the Catholic Church beginning in the 13th century (Baumer, 1977; Skinner, 1979b; Tannenbaum & Schultz, 2004).

Up until the 13th century, the writings of St. Augustine had dominated the Church. The once secular writings of the ancient Greeks had all but disappeared from Christian Europe, surviving in the Arabic and eventually Islamic worlds. But translations of Aristotle's works prepared by Bishop Raymond of Toledo and William of Moerbeke began to emerge in Spain in the early 13th century (Skinner, 1979b).

These writings discussed Aristotle's more secular views on government and his visions of human nature that contrasted with those of Christianity. St. Thomas Aquinas attempted to reconcile Aristotle with Christianity, as did other writers, such as Dante Alighieri (1265–1321) and then Marsilius of Padua (ca. 1275–1343). In various ways, their importation of Aristotle into Christianity often left room for a political dualism that distinguished secular from spiritual rule and allowed for potential limits on the authority of kings (Sabine & Thorson, 1973; Skinner, 1979b).

The growing importation and influence of Aristotle challenged Christian doctrine. Luther and Calvin doctrinally disagreed with the new direction of the Church. For different reasons they challenged the pope's reading of the Bible, and they sought political sanction and support from local rulers in Germany to protect themselves (Tannenbaum & Schultz, 2004). Luther also disagreed with many practices of the Church, such as the selling of religious indulgences, that is, the forgiveness of sins, for a price. Additionally, kings, whose power was growing, chafed under the popes. Some, such as Henry VIII of England, fought bitterly with the Pope to secure a divorce. Eventually, Luther, Calvin, and many kings came to challenge a key assertion of the Christian world order: papal infallibility (Skinner, 1979b). The Ninety-Five Theses that Luther nailed to the church door in Wittenberg in 1517 encapsulated his grievances with the pope. Finally, Luther's challenge to papal infallibility questioned the belief that there was only one correct interpretation of the Bible. Instead of the Church's having one ruler, Luther described the Church as an institution or community of equality. It was of a priesthood of all believers in a world where, in the view of St. Augustine, all were equally sinners. Similarly, John Calvin, in *Institutes of the Christian Religion* (1535), sought to define the essence of Christian liberty as a spiritual force, thereby seeking to distinguish civil and religious institutions in terms of the scope of their authority (Sabine & Thorson, 1973; Skinner, 1979b; Tannenbaum & Schultz, 2004). Yet despite distinct realms, both types of institutions aimed to help humans serve God. Monarchs thus received directly from God their authority to rule.

What emerged politically from the Protestant Reformation was the first crack in the unified world order of medieval Christianity. There was no longer one uninterrupted chain of being, from God to popes to kings to men. There was at least a dualism in faith. What emerged from this criticism of the pope was that kings were significantly invested with new power and authority. Instead of having to seek their authority to rule from the Church and the pope, increasingly monarchs claimed a *divine right* to rule, in other words, political legitimacy directly from God. Monarchs thus had even more authority than before. God directly sanctioned their unlimited power to act and the duty of subjects to obey them. Although the average subject might not have noticed much difference from all-powerful monarchs claiming their power from the pope as opposed to God directly, the Protestant Reformation

could be seen as opening up space for the justification of independent secular institutions to rule.

The Emergence of the Enlightenment and the Copernican Scientific Revolution

The medieval political order was supported by a theology and a cosmology. The theology was supplied by Christianity, the cosmology by Ptolemy (Koyré, 1982; Kuhn, 1992; Lovejoy, 1976). Ptolemy was a 2nd-century Roman astronomer whose book *Almagest* constructed a model of the universe that depicted the earth at the center, with the sun, the moon, the planets, and the stars all revolving around it. This geocentric model provided a cosmological model that supported the Christian Church in that depicting the earth as the center of the universe meant that God had chosen earth, Christianity, and ultimately the pope to be his direct descendants in terms of exercising authority. In effect, the pope was the head of a religion at the center of the universe that governed over all the earth. God was thus wrapped into the very fabric of the universe.

Over time, the geocentric model proved increasingly difficult to sustain. Astronomical instruments and mathematical models failed to predict adequately the actual movement of the planets and stars across the heavens. There was a disconnect between what Ptolemy's model predicted and what was observed. It thus became necessary to make the model increasingly more complex in order to describe accurately what was happening in the sky. Eventually the model proved clunky and difficult to maintain (Kuhn, 1992).

Nicholas Copernicus

Nicholas Copernicus (1473–1543) was a Polish astronomer who sought to address the problems with Ptolemy's model. In his 1543 book *On the Revolution of the Celestial Spheres*, Copernicus attempted to correct the errors of the geocentric model by proposing instead that the sun and not the earth is at the center of the universe. With such a proposal, many of the problems of the older geocentric model were solved, and it was easier to predict the orbits in the sky. One would not think that simply proposing a new astronomical theory would cause a political crisis, but it did. Copernicus was afraid to have his book published until after his death, fearing excommunication from an angry Church. In 1616, the pope declared the book heresy and barred Catholics from reading it (Blumenberg, 1987a; Koyré, 1982; Kuhn, 1992; Tannenbaum & Schultz, 2004).

Signs that the Church would not accept an alternative model of the universe lightly or willingly were confirmed with the condemnation of Galileo (1564–1642). It was his observations in the early 17th century of the moons orbiting Jupiter that led him to believe that Copernicus was correct. His run-ins with the Church demonstrated the threat the newly emerging sciences posed to religion, and they

were part of the broader movement during the 17th century, known as *the Enlightenment* (Kuhn, 1992).

Origins of the Enlightenment

The Enlightenment was a major intellectual challenge to the medieval order (Blumemberg, 1987a; Tannenbaum & Schultz, 2004). Immanuel Kant (1724–1804), a late-18th-century philosopher, once wrote an essay titled “What Is Enlightenment?” by saying that its motto was “Dare to know!” (Kant, 1979, p. 54). The Enlightenment was a rejection of accepting truths based on faith; instead it was an appeal to reason and eventually to experience or empiricism to gather knowledge. The Enlightenment was also an appeal to use reason as a tool to gather political knowledge, and to ascertain truth (Baumer, 1977). The Enlightenment was driven first by the reemergence of classical Greek philosophy and eventually by the adoption of new scientific methods of gathering knowledge. Thus, in 1620, Francis Bacon (1561–1626) argued in *Novum Organum* that one must proceed from new foundations to achieve new certainty of knowledge. Bacon stressed a new method of inquiry, eventually called the scientific method, for studying the world. René Descartes’s (1596–1650) *Meditations on First Philosophy* (1641) sought to use doubt as a process of arriving at truth by seeking to doubt all until doubt was no longer possible. Descartes’s self-doubting took him back to the question of his own existence. At this point, he arrived at a certainty of knowledge with the exclamation *cogito ergo sum*, or “I think, therefore I am.” Isaac Newton (1642–1727) developed the basic laws of motion, gravity, and physics in his *Principia Mathematica* (1687), which portrayed a universe that could be explained by math and science and not by God.

Taken together, Copernicus, Galileo, Bacon, Descartes, and Newton provided a scientific approach to knowledge that was not contingent on God (Tannenbaum & Schultz, 2004). Their works further cracked the unity of the medieval theology and cosmology by evicting the earth from the center of the universe and replacing old knowledge with knowledge based on new rational and scientific foundations. The legacy of these Enlightenment thinkers provided a new way for political thinkers to talk about the world. One could use the tools of reason to challenge the old political order and think about a political order that was secular and rational.

Liberalism and Social Contract Tradition

The Origins of the Social Contract Tradition

The Protestant Reformation cracked the singularity of the Christian world but left monarchs in a stronger position than before by investing them with divine right to rule. The Copernican scientific revolution challenged old truths by questioning religious knowledge

and truth. Both movements were threats to the medieval order, but they also were in conflict with one another. Inevitably they would collide, and they did so when it came to the power of princes and the duty of subjects to obey them.

The Protestant Reformation gave monarchs a religious justification to rule, yet a persistent problem surrounded the scope of their authority (Skinner, 1979a, 1979b). The divine right of royal power meant that monarchs had unlimited power, but what if it were abused and kings turned out to be tyrants? One early sign of efforts to trim the excesses of kings occurred at Runnymede, England, in 1215, when noblemen forced King John to sign the Magna Carta and recognize some basic rights of the people he ruled.

The Protestant Reformation also sowed some of the seeds of limitation of royal power. While divine right generally meant absolute obedience to the ruler, at least in Protestant states, in countries such as Catholic France, religious minorities experienced religious persecution. French Protestants, known as Huguenots, began to construct constitutional theories justifying some passive disobedience to rulers who exceeded their powers. Francis Hotman’s 1573 *Franco Gallia* was one of the first tracts to explore this subject. Then in 1579, the anonymously authored *Vindiciae Contra Tyrannos* forcefully developed the case against tyrants who abused their authority, at least in terms of matters of faith (Skinner, 1979b). The book addresses four questions: whether subjects are obligated to obey rulers if the rulers command anything against God, what means subjects may use to resist, how far subjects may go to resist, and whether neighboring princes can aid subjects in resisting. *Vindiciae* asserted that monarchs had a duty to support the pure doctrine of God, and should they fail, subjects had a religious duty to resist. In effect, rulers who defied God broke a contract connecting them to God and to their subjects.

Other French thinkers took a different tack to address religious persecution and the power of monarchs. Jean Bodin’s (1530–1596) *Six Books of the Commonwealth* made the case for religious toleration and the separation of the Church from the state (Keohane, 1980; Sabine & Thorson, 1973). Bodin also developed another significant argument, claiming that sovereignty, or political authority in a society, originated not with the monarch but with the people. The significance of his arguments on toleration and sovereignty was to suggest that kings were limited in their power by the people and that matters of state did not extend to those of faith. Perhaps in recognition of these arguments, in 1598 King Phillip of France issued the Edict of Nantes, which granted Protestants freedom of conscience.

England and the Battle Between the King and Parliament

In the 16th and 17th centuries in England, justifications of royal power took place against the backdrop of political battles between the king and a restless parliament. In 1628,

the Petition of Right was adopted to limit the royal prerogatives of King Charles I. Then in 1649, Charles, an advocate of the divine right theory, was ousted by Oliver Cromwell, a member of the British Parliament. England was declared to be a commonwealth until the monarchy was restored in 1660. In 1688, the Glorious Revolution led to the removal of Catholic James II and the adoption of the English Bill of Rights in order to protect individual rights against monarchical excesses.

Thomas Hobbes

The battle for political power, however, took place not simply in parliament but also in books. Thomas Hobbes (1588–1679) was a royalist, a defender of monarchical power. His most famous book, *Leviathan* (1651), sought to address the disorder caused by the attacks on the monarchy and brought together many of the current intellectual trends. The *Leviathan* draws on the emerging science and rationalism of the Enlightenment to forge a new theory of royal absolutism premised, not on divine right, but on a different argument that appealed to what has come to be called *social contract theory* (Hampton, 1988; Rawls, 2007; Riley, 1982; Strauss, 1952; Wolin, 1960).

A basic assumption of politics dating back to the ancient Greeks was that politics and the state were natural (Strauss, 1953). In other words, both had always existed, and there was never a time when a presocial situation existed, one with no government or political authority. The significance of this argument, especially as it evolved in the medieval Christian world, was that if government was natural and had always existed, then it was simply natural that people had to obey the law and their rulers, especially if the latter were sanctioned to rule by God. Political sovereignty, to use Bodin's term, was situated with the rulers and not the people. Monarchs were sovereign; they had political power, not the people, and therefore the people had no ability to limit an abusive ruler. But if, as Bodin suggested, sovereignty belonged to the people, then monarchs received their power from God indirectly, through the people, so long as the people did not abuse that power. This at least was the argument in the *Vindiciae*. The *Vindiciae* envisioned political society as similar to a contract that linked the people to the king and to God. If the king broke the contract, then the people had a right to remedy that breach.

Thomas Hobbes and other political writers extended this contract metaphor. They sought to explain the origin of political power by imagining a distant past when no government existed at all (Pateman, 1988; Rawls, 2007; Strauss, 1952, 1953; Tannenbaum & Schultz, 2004; Wolin, 1960). In *Leviathan*, Hobbes (1651/1998) imagined a *state of nature*, a presocial situation, well before government or civil society existed. In that state of nature, individuals existed as isolated creatures. But the creatures Hobbes described were like machines in motion, seeking to stay in

motion. Moreover, Hobbes ascribes to humans a basic psychological impetus: “a general inclination of all mankind, a perpetual and restless desire of power after power that ceases only in death” (p. 161). In this state of nature, all humans have one basic right, a right to self-preservation and to pursue power. But in pursuing both, conflicts emerge, especially because there is no government to keep people in line. For Hobbes, the state of nature is thus a state of war, in which life is “solitary, poor, nasty, brutish, and short” (p. 100).

Hobbes explained that the desire for self-preservation leads individuals to forge a social contract. This is a contract among individuals to turn power over to some individual to exercise absolute power to enforce laws and keep order. Individuals in a state of nature thus relinquish, by way of their consent and a social contract, power to an absolute ruler to keep the peace. Government is thus an artificial machine or entity given the power to keep the peace. It is no more than simply the sum of the individuals who make it up. This image is aptly captured in the original book cover for *Leviathan*, which depicts a monarch drawn in such a way that he is composed out of the sketches of a multitude of individuals (Sabine & Thorson, 1973; Tannenbaum & Schultz, 2004; Wolin, 1960).

The power of *Leviathan* resides in its argumentative force and logic (Strauss, 1952; Tannenbaum & Schultz, 2004). It is an account of royal power devoid of religion and appeals to divine right. Instead, it draws on rationalism and the empiricism and science of the Enlightenment to construct a theory of human nature, society, and government that does not rely on any references to God. It is a mechanistic view of humans and society, much like the vision of the universe described by Newton's laws of motion. It appropriates the newly emerged social contract tradition to defend an old proposition, royal absolutism (Wolin, 1960).

John Locke and the Rise of Liberalism

But not all defenses of the monarchy veered in the direction that Hobbes took. Others relied on traditional appeals to God and the Bible to assert royal prerogative. One famous effort to do that was Sir Robert Filmer's 1680 *Patriarcha* (Ashcraft, 1986). Several traditional metaphors were often used to justify political power, especially monarchical authority. One claim was the idea that kings were like fathers (Schochet, 1975). By that, if fathers were the heads of families, governments could be viewed as large families over which kings ruled. Moreover, in the 17th century (and earlier), fathers in families were viewed as having unlimited power over their wives and children. Thus, for Filmer, if kings were like fathers and they had absolute power over the family, the same would be true with kings over governments. Second, Filmer defended monarchical power with biblical references, often by invoking the Fifth Commandment—“Honor thy father and

mother”—to support monarchical rule (Schochet, 1975). The book of Genesis was also invoked to defend royal power. Genesis describes Adam as given dominion over the world, and that grant of authority was interpreted to mean that government was natural and that Adam and his descendants had absolute power over all. Overall, Filmer's *Patriarcha* combined biblical, familial, and other metaphors to defend royal absolutism.

Perhaps the most famous effort to refute Filmer and to defend limited monarchy is found in the writings of British political philosopher John Locke (1632–1704). His 1690 *Two Treatises on Government* is a response to Filmer that refutes claims for absolute monarchical power and justifies both parliamentary authority and a revolution to limit the authority of the crown (Cranston, 1957; Skinner, 1979b). To undertake this task, Locke also invoked social contract reasoning or metaphors to make an argument about government and society. Like many of his contemporaries, Locke sought to explain the origin of society. This included discussing why laws exist, why they should be obeyed, and who should be entitled to rule a country.

But Locke was writing at a time when the power of kings was being questioned and parliaments and popular governments were beginning to develop. Like Hobbes, Locke began his claims about government by assuming a presocial state of nature (Rawls, 2007; Riley, 1982). This state of nature is a condition of natural liberty in which all are free and equal. No human laws or rules exist in this state of nature. This does not mean that there are no rules of justice. Locke talks about the existence of natural laws or rule of justice. More important, he indicates that individuals in the state of nature possess certain *natural rights*. These rights include the right to defend oneself, to defend other rights, and to act in ways to preserve oneself and one's possessions. But the most important natural rights for Locke, in the *Two Treatises*, draw linkages between property, liberty, and government (Ashcraft, 1986; Cranston, 1957; Macpherson, 1962). Although Locke describes a state of nature as one of perfect freedom and a place where individuals have natural rights, these rights are unclear. Moreover, while in the state of nature, one may come to acquire possession of items; again, the exact rights to these objects are occasionally uncertain. Others may try to take them. Left unchecked, the state of nature can turn into anarchy or a state of war. Individuals enter in a social contract to protect their natural rights. The social contract, civil society, or government gives clarity to natural rights, including the right to property. The preservation of property rights is the chief goal of civil society.

Property for Locke refers to one's life, liberty, and estate. The goal of the social contract in creating civil society is protecting property (Macpherson, 1962). The law gives legal meaning, status, and protection to property. Thus, on one level, the social contract is one that seeks to protect preexisting natural rights. These are rights that all individuals possess. In creating a government or civil

society, individuals reach an agreement among themselves to create an authority that will protect and enforce their rights. The contract is not between the king or the government and the people (Pateman, 1988; Skinner, 1979b; Tannenbaum & Schultz, 2004; Wolin, 1960). The contract is among the people and it defines the scope of powers and authority of the government. The people contract among themselves and literally give a contract to the government to act on their behalf.

In structuring a government to protect the rights of the people, several points are important to stress. The first is the concept of *consent*. Whereas in a monarchy, in which people do not choose to be ruled or, rather, do not give consent to a king to rule over them, government, for Locke, is all about consent (Riley, 1982). That is, the authority to act and, with that, the duty to obey the law are premised on the idea of consent. But does this concept of consent mean an active, explicit form of consent on every major decision made by the government? The answer Locke provides is no. Here he invokes two concepts, the idea of *tacit consent* and that of *majority rule*, arguing that entrance into civil society and therefore the obligation to obey the law are not based on an expressed consent. One can tacitly consent to obeying the law and entering civil society. Tacit consent may be premised on the benefits one enjoys in society. In effect, the very fact that one benefits from the law means that one is consenting, at least tacitly, to the legitimacy of the rules.

Does consent mean that for government to be legitimate, the people have to agree all the time with everything the government undertakes? Again Locke's answer is no, and here he introduces the concept of majority rule, whereby assemblies or parliaments may vote by majorities to act on behalf of the people. The concept of majority rule is crucial for Locke (Ashcraft, 1986). Once an initial social contract is formed among the people and the government is instituted, unanimous consent (or at least express consent) is not required. Instead, society may act if a majority of the population or parties to the contract (or the government) supports an idea. Government and civil society then are premised on a notion of majority consent or expression. In essence, this is the concept of a popular government in which the people are ruled through majorities and not on the whim of one individual (king) or a minority (aristocracy).

With the joining of tacit consent and majority rule, the question of minority rights arises. Could not a majority of the population vote to enslave the minority or to strip them of their property rights? This is the classic problem of a popular government. Locke has several solutions to preventing this tyranny. First, recall that the rights to property are natural, meaning that these natural rights; operate as outside limits on what the government can do. Government is instituted so that it can protect these natural rights; one cannot create a government that would seek to deny or suppress these rights. Second, the entire concept of the

social contract is meant to address this problem. Specifically, one would presume that no individual would consent to creating a civil society that would suppress or limit rights. Instead, the very notion of the contract and consent should also operate as a mechanism that checks abuses. But finally, there is a third mechanism to limit abuses, the notion that government or government officials are trustees for the people, seeking to act in the people's best interest.

In order to make the argument that government serves as trustees for the people, Locke draws some parallels between government and parents. Recall that monarchies were defended at this time by contending that kings were like fathers who have absolute rule over their wives and children. The family was thus viewed as a minimonarchy. Locke challenges that notion of the family. First, Locke argues that parental power over children is not absolute and natural but instead is a guardian relationship. By that, he meant that parents cannot do what they want with their children. Parents' powers over children are limited, with parents' serving as a "temporary government" over their children until the children reach adulthood. Locke also redefines the marriage relationship, describing it as voluntary. The family, then, is not natural but a product of convention (contract), with the powers of the father limited to looking out for the best interests of the wife and children (Schochet, 1975).

If the father's power is limited to that of being a guardian, and if the family is the metaphor for the government, then the role of political leaders is one of a limited, guardian relationship. In effect, government is to serve the public good and protect property (life, liberty, and estate). What if government breaches that trust? Near the end of the *Two Treatises*, Locke reserves to the people the right to decide when the government has abused its trust and therefore when they should change the government. In effect, Locke describes a right not only for individuals to create a civil society and government but also a right to revolt to dissolve it. The people decide when it is time to change government (Skinner, 1979b).

Emerging out of Locke's writing is what has been called a political theory of *liberalism* (DeRuggerio, 1964). This theory of liberalism should not be confused with the notions of conservative or liberal that are popularly invoked in politics and the press today. Instead, what is called Lockean liberalism is a set of values committed to individual rights, limited government, and the belief that individuals are sovereign, that is, that they can create or end a government, whether by elections or more extraordinary means. Emerging out of Locke's *Two Treatises* are ideas such as consent of the governed, majority rule, respect for minority rights, and the ideas that government officials are trustees looking out for citizens and serving the public good. Locke's rethinking of government based on consent shifts the notion of political authority away from justifications resting on religious or biblical authority.

In the *Two Treatises* as well as in his other writings, such as *A Letter Concerning Toleration* (1679), Locke also endorses religious toleration, further strengthening the notion of a separation of Church and state.

Jean-Jacques Rousseau

Locke was not the only writer to invoke a social contract theory for describing the origins of government. Jean-Jacques Rousseau (1712–1778) was a social contract theorist who tried to explain the origin of society by appealing to an ancient compact reached among individuals (Riley, 1982; Tannenbaum & Schultz, 2004). This compact created the institution of property, but in contrast with Locke, property for Rousseau was not depicted in a positive light. Instead, property, along with all of society and government, were described in his *Discourse on the Origins of Inequality* (1755/1964) as a trick by the rich and powerful over the poor. Rousseau expressed this idea as follows:

The first man who, having enclosed off a piece of land, got the idea of saying "*This is mine*" and found people simple enough to believe him was the true founder of civil society. What crimes, what wars, what murders, what miseries and horrors would someone have spared the human race who, pulling out the stakes or filling in the ditch, had cried out to his fellows, "Stop listening to this impostor." (Second Discourse, p. 141)

For Rousseau, the original social contract and the property that it created were the first step in the gradual and eventual enslavement of individuals. Elsewhere in the Second Discourse, Rousseau sees the first step in the creation of property as leading to even further social conflicts and distinctions, such as family inequalities. The question for Rousseau was thus how to take humans off this path, which started with a fall from their state of natural equality and led them into the inequalities they experience in civil society. As Rousseau (1762/1977) would state later, in his most famous book, *The Social Contract*, "Man is born free, but he is everywhere in chains" (p. 49). Although in nature humans were free and equal, social institutions such as private property take away this natural freedom. Rousseau argued in *The Social Contract* that a genuine contract among the people is the only legitimate basis for forming a government (Rawls, 2007). Yet, he states, government still faces a basic problem:

How to find a form of association which will defend the person and the goods of each member with the collective force of all, and under which each individual, while uniting himself with others, obeys no one but himself, and remains as free as before. (p. 60)

The primary task of *The Social Contract* is reconciling government with individual freedom (Sabine & Thorson, 1973; Shklar, 1969). Rousseau accomplishes it with a social

contract that gives up natural physical freedom for moral or civil autonomy. Individuals are free when they live according to laws they have consented to or constructed for themselves. They achieve freedom when, according to Rousseau, their individual freedom conforms to the general will. All citizens are expected to obey the general will, the laws made by a lawgiver chosen by the people. Rousseau's arguments thus appear to declare a very broad notion of democracy that entrusts to the people the authority to rule and make laws for a society. But what if the people err? Can individuals disobey the law? The Rousseau of the Second Discourse calls for revolution or rebellion if the contract is a false one premised on fraud, but the Rousseau of *The Social Contract* states that individuals may be "forced to be free," or be compelled to follow the law in order to promote the type of government and freedom he advocates.

Liberalism and the Social Contract Tradition

The social contract tradition, especially as articulated by Locke and Rousseau, strongly criticized royal power. Together they placed the origins of political power in the people. It was through a contract that the people expressed their sovereignty, their authority to define the legitimate ends of government. The notion of a social contract is at the root of concepts such as constitutionalism, limited government, majority rule, and respect for individual rights. The social contract tradition, especially through Locke, formed the basis of what has come to be called the *liberal tradition* (Rawls, 2007; Riley, 1982). Liberalism, expressing this commitment to individual rights and a limited a popular government, is one of the major theories or philosophies of modern political thought (DeRuggerio, 1964). Liberal social contract theory has been an important tool for many political thinkers, including those of the present, to defend theories of political authority that respect rights. In the 18th century, another notable liberal, Mary Wollstonecraft (1759–1797), appealed to liberal social contract theory in *A Vindication of the Rights of Women* (1792) to extend its claims to women (Sapiro, 1992; Tannenbaum & Schultz, 2004). Thomas Paine (1737–1809) drew on liberal values, both in his 1776 *Common Sense* and in his 1791 *Rights of Man*. The former called for an American revolution from England, and the latter defended the French Revolution of 1789.

But not everyone in the 18th century or even today agrees with all the values of liberalism and the use of a social contract to define political authority. David Hume (1711–1776) was an important British philosopher best noted for his *Treatise of Human Nature* (1739), a major book on epistemology and knowledge. He also penned numerous political essays, some of which criticized the social contract tradition. He argued that contracts were not the basis of political authority. Instead, mere habit or custom gave rise to government. But Hume did concur that rules of

justice and authority were conventional, not natural, and thus he sided with many claims used traditionally to support political authority. Additionally, some contemporary scholars such as Carol Pateman (1988) and Charles Mills (1997) have argued that the social contract tradition was exclusive, that is, it excluded women and people of color from the social contract, and thereby it created a sexual and racial contract that defended sexism and racism. Finally, Okin (1979) and Lloyd (1984) see a male-centered or misogynist bias in the employment of Enlightenment reasoning.

The French Revolution

England was not the only country to oust kings. The French Revolution of 1789 was the fourth major revolution to help define modernity, and it was launched in the spirit of Rousseau, liberalism, and the social contract tradition.

The French Revolution is considered one of the most significant political events in modern European or Western history (Baumer, 1977; Horkheimer & Adorno, 1972). The mass demonstration beginning the French Revolution commenced on July 14, 1789, now known as Bastille Day. This was the day that the Bastille Prison, used by the king to jail political and other prisoners, was attacked by the people. The prison was a symbol of royal abuse. With the chant of "Liberty, equality, and fraternity" as its motto, the French Revolution deposed, at least temporarily, the monarchy of France and beheaded both King Louis XVI and Queen Marie Antoinette. In its place was instituted a people's government with a parliament-like body that made laws for the country.

The French Revolution was significant in at least a couple of ways. First, it was the dismantling of a powerful monarchy by the people. It was in many ways the capstone, or coming together, of many of the ideas that had been articulated about state and society beginning with the *Vindiciae* if not earlier. It was a secular revolution of the people, expressing the view that royal power could be limited and ultimately dethroned by the people in the name of their own rights (Keohane, 1980). Second, the French Revolution, coming just a few years after the American revolutionary war, appeared to set the tone for a rethinking of political authority in Europe. It foreshadowed the end of absolute monarchies, divine right, and the rise of constitutionalism and individual rights. Eventually, by 1848, revolutions by peoples across Europe would oust other monarchies, paving the way for modern democratic societies governed by parliamentary bodies that eventually would be elected on the basis of universal suffrage.

But not all viewed the French Revolution favorably. Irishman Edmund Burke (1729–1797) criticized it in his *Reflections on the Revolution in France* (1790). He argued that the revolution undertaken in the name of reason, Rousseau, and the universal rights of men was destructive.

Unlike the American revolution, which he supported as a defensive effort to protect rights, he considered the French Revolution an example of the destructive power of abstract reason gone wild. He rejected claims that the social contract was the origin of society, seeing instead the forces of history and tradition as the basis of what holds a society together. A true social contract holds the past, present, and future together. The French Revolution had destroyed that contract, revealing a host of problems. According to Burke, the error of the traditional contract theorists was in believing that all individuals were equal and that they had rights that included the right of revolution in order to create and express their other rights. In contrast with Locke and Rousseau, Burke maintained that a natural aristocracy existed that was best suited to rule and that the best way to protect rights was generally to obey tradition and authority (Sabine & Thorson, 1973).

The Kantian Revolution

The last revolution to frame the Enlightenment and early modern political thought is the Kantian one, named after the German (Prussian)-born philosopher Immanuel Kant (1724–1804; Blumenberg, 1987a; Kuhn, 1992; Tannenbaum & Schultz, 2004). Kant's revolution is one of reason and rationalism. It is a coming together of several political and philosophical positions into a theory that both encapsulated and transformed modernity. Kant is probably best known for his *Critique of Pure Reason* (1787/1965), a major book on philosophy and knowledge, but he also wrote on moral, ethical, and political philosophy.

Kant's primary contribution is in philosophy. He sought to overcome a gap between two theories of knowledge that had emerged during the Enlightenment as alternatives to faith (Tannenbaum & Schultz, 2004). One theory was that all knowledge was based simply on reason, and the other theory that all knowledge was empirical or based on the senses. Previous writers such as Descartes thought logic or rationalism was the font of all knowledge. Locke and Hume were empiricists who argued that all knowledge was based simply on ideas derived from experiences of the world. The problem with this theory was that no one could prove empiricism correct—how could one empirically show that ideas in our head corresponded to objects that exist around us? Conversely, pure rationalism could not be proved correct. How could mere logic or reason give humans knowledge of the outside world? Kant (1787/1965) invokes Copernicus's switch from a geocentric to a heliocentric model of the universe as a way to address this problem:

Hitherto it has been assumed that all our knowledge must conform to objects. But all attempts to extend our knowledge of objects by establishing something in regard to them a priori, by means of concepts, have, on this assumption, ended in failure. We must therefore make trial whether we may not

have more success in the tasks of metaphysics, if we suppose that the objects must conform to our knowledge. . . . We should then be proceeding precisely on the lines of Copernicus' primary hypothesis. (p. 22)

Using the Copernican analogy, Kant argues that our empirical experiences of the world are filtered or conform to categories of understanding that are innate. In effect, there are limits to human understanding.

The significance of Kant's arguments is that there are certain things that cannot be proved, such as the existence of God. Thus, Kant's arguments dramatically question any ability to invoke religion or theology to sustain theories of political obligation or authority. Kant's arguments render assertions about natural law and an ultimate cosmology as grounds for politics impossible to defend, thereby finally killing off the great chain of being, from God to popes to kings to men. Moreover, if people wish to preserve political and philosophical truths, they may have to assume they are true or act "as if" they are true because it may be impossible to prove their validity by simple appeals to reason (Blumenberg, 1987a).

Kant's arguments have other political implications. He rejects the ideas that a historical social contract ever existed and that it was the basis of political obligation (Rawls, 2007; Riley, 1982). He believed, however, that the idea of the social contract, although not a historical reality, serves as a regulative idea for the commonwealth, an ideal to which the constitution and all laws must aspire. It, along with human freedom, are ideas that act as both goals and assumptions defining human political society. Moreover, Kant's social contract, as well as his moral philosophy, draws heavily on Rousseau. Freedom, or autonomy, is defined as individuals' acting according to laws that they will for themselves. This concept of moral freedom serves as the basis of a constitutional government that protects individual rights.

The significance of Kant's revolution is that it not only used reason to question old truths (an important characteristic Kant shares with other Enlightenment writers) but also showed the limits of the new ways of knowing. Kant's writings reveal the powerful force of modernity: It both enabled a capacity to be skeptical of and question old political truths, dogmas, and authority and provided the building blocks for a new order, a redefined political authority premised not on divine right and religion but on individual rights and limited government.

Conclusion

Friedrich Nietzsche (1844/1967), a 19th-century German philosopher, once wrote, "Since Copernicus, man seems to have gotten himself on an inclined plane—now he is slipping faster and faster into—what? Into nothingness?" (p. 155). This statement well captures the sentiment and import of early modern political thought. Five revolutions

taking place between 1500 and 1800 dismantled the ancient-medieval world of Christian unity. It destroyed a political theory of an essentially absolutist Christian state, re-creating a new one premised on a secular vision of a constitutional government committed to individual rights. This is the core vision of political liberalism.

The project of the Enlightenment or modernity would continue to develop and build on liberalism in the 19th century. For example, political thinkers such as Jeremy Bentham, James and John Stuart Mill, and Benjamin Constant, among others, would transform liberalism into the moral economic theory of utilitarianism (Halevy, 1955), and Karl Marx and his followers would push the theory toward socialism. These theories and others were indebted to the arguments that the thinkers of the Enlightenment produced.

References and Further Readings

- Ashcraft, R. (1986). *Revolutionary politics & Locke's two treatises of government*. Princeton, NJ: Princeton University Press.
- Baumer, F. L. (1977). *Modern European thought: Continuity and change in ideas, 1600-1950*. New York: Macmillan.
- Blumenberg, H. (1987a). *The genesis of the Copernican world*. Cambridge: MIT Press.
- Blumenberg, H. (1987b). *The legitimacy of the modern age*. Cambridge: MIT Press.
- Cranston, M. (1957). *John Locke: A biography*. London: Longman.
- DeRuggerio, G. (1964). *The history of European liberalism*. Boston: Beacon Press.
- Halevy, E. (1955). *The growth of philosophical radicalism*. Boston: Beacon Press.
- Hampton, J. (1988). *Hobbes and the social contract tradition*. New York: Cambridge University Press.
- Hobbes, T. (1988). *Leviathan* (C. B. Macpherson, Ed.). New York: Penguin Books. (Original work published 1651)
- Horkheimer, M., & Adorno, T. W. (1972). *The dialectic of the Enlightenment*. New York: Continuum.
- Kant, I. (1965). *A critique of pure reason* (N. K. Smith, Trans.). New York: St. Martin's Press. (Original work published 1787)
- Kant, I. (1979). An answer to the question: What is enlightenment? In H. Reiss (Ed.), *Kant's political writings* (pp. 54-61). Cambridge, UK: Cambridge University Press.
- Keohane, N. O. (1980). *Philosophy and the state in France: The Renaissance to the Enlightenment*. Princeton, NJ: Princeton University Press.
- Koyré, A. (1982). *From the closed world to the infinite universe*. Baltimore: Johns Hopkins University Press.
- Kuhn, T. S. (1992). *The Copernican revolution: Planetary astronomy in the development of Western thought*. Cambridge, MA: Harvard University Press.
- Lloyd, G. (1984). *The man of reason: "Male" and "female" in Western philosophy*. Minneapolis: University of Minnesota Press.
- Lovejoy, A. O. (1976). *The great chain of being: A study of the history of an idea*. Cambridge, MA: Harvard University Press.
- MacIntyre, A. (1984). *After virtue*. Notre Dame, IN: University of Notre Dame Press.
- Macpherson, C. B. (1962). *The political theory of possessive individualism: Hobbes to Locke*. Oxford, UK: Clarendon.
- Masters, R. D. (1968). *The political philosophy of Rousseau*. Princeton, NJ: Princeton University Press.
- Mills, C. W. (1997). *The racial contract*. Ithaca, NY: Cornell University Press.
- Nietzsche, F. (1967). *On the genealogy of morals*. New York: Vintage Books. (Original work published 1887)
- Okin, S. M. (1979). *Women in Western political thought*. Princeton, NJ: Princeton University Press.
- Pateman, C. (1988). *The sexual contract*. Stanford, CA: Stanford University Press.
- Rawls, J. (2007). *Lectures on the history of political philosophy*. Cambridge, MA: Belknap Press.
- Riley, P. (1982). *Will and political legitimacy: A critical exposition of social contract theory in Hobbes, Locke, Rousseau, Kant, and Hegel*. Cambridge, MA: Harvard University Press.
- Rousseau, J. J. (1964). *The first and second discourses* (R. D. Masters, Ed.). New York: St. Martin's Press. (Original work published 1755)
- Rousseau, J. J. (1977). *The social contract* (M. Cranston, Ed.). New York: Penguin Books. (Original work published 1762)
- Sabine, G. H., & Thorson, T. L. (1973). *A history of political theory*. Hinsdale, IL: Dryden Press.
- Sapiro, V. (1992). *A vindication of political virtue: The political theory of Mary Wollstonecraft*. Chicago: University of Chicago Press.
- Schochet, G. J. (1975). *Patriarchalism in political thought*. New York: Basic Books.
- Shklar, J. (1969). *Men and citizens: A study of Rousseau's social theory*. New York: Cambridge University Press.
- Skinner, Q. (1979a). *The foundations of modern political thought: Vol. 1. The Renaissance*. New York: Cambridge University Press.
- Skinner, Q. (1979b). *The foundations of modern political thought: Vol. 2. The age of Reformation*. New York: Cambridge University Press.
- Strauss, L. (1952). *The political philosophy of Hobbes: Its basis and its genesis*. Chicago: University of Chicago Press.
- Strauss, L. (1953). *Natural right and history*. Chicago: University of Chicago Press.
- Tannenbaum, D., & Schultz, D. (2004). *Inventors of ideas: An introduction to Western political thought*. Belmont, CA: Wadsworth.
- Wolin, S. (1960). *Politics and vision: Continuity and innovation in Western political thought*. Boston: Little, Brown.

70

NEOCLASSICAL LIBERALS

ELLEN GRIGSBY

University of New Mexico

In the 19th century, liberals responded to the changing economic and social conditions associated with industrialization and urbanization in a variety of ways. One group of liberals argued for the continuing relevance of classical liberalism's emphasis on limited government whereas another group of liberals began to make the case for an expansive role for government. Although both groups championed the importance of individual liberty (and thus firmly insisted on their own philosophical legitimacy as the "true" liberals), the two groups increasingly diverged on such matters as the role of the state in promoting liberty, the optimal reach of government intervention in the economy, and the appropriate governmental response to social problems such as poverty. The "new" classical liberals, that is, the *neoclassical* liberals, argued that the liberalism that Locke had envisioned in his 1690 *Second Treatise*, as they interpreted it, continued to be the best guide for evaluating and directing political practice, but the second group—which came to be known as welfare liberals and which included writers such as T. H. Green (1836–1882)—called for a more broadly regulative government than Locke's limited "umpire" state (see Ball & Dagger, 2009; Kramnick, 1998).

This chapter discusses the defining elements of neoclassical liberalism—advocacy of (a) limited government and (b) maximization of individualism or individual liberty, with individualism or liberty being defined as that

which is realized when self-interested persons make decisions for themselves in the absence of governmental interference. A major focus of the discussion is the political writings of British theorist Herbert Spencer (1820–1903) and American theorist William Graham Sumner (1840–1910). Spencer and Sumner were leading neoclassical liberals of the 19th century, and a close examination of their writings reveals a complex interplay of 17th-century Lockean ideals with 19th-century economic and intellectual developments.

Herbert Spencer

British philosopher and sociologist Herbert Spencer has often been called both a neoclassical liberal and a social Darwinist. The latter designation is somewhat problematic, not because Spencer did not think in evolutionary terms, but because he began applying evolutionary perspectives in his writings before Darwin published his own findings. Indeed, Spencer, rather than Darwin, first employed the phrase "survival of the fittest" (Offer, 1994, xvii–xviii).

In *The Proper Sphere of Government*, Spencer (1843/1994d) argued that individuals have natural rights to life and property. With respect to the latter, Spencer, like Locke, looked to both religious and historical sources to support his claim for the naturalness of private property

and for its existence as a “necessary means” of meeting human needs. By “natural,” Spencer, as Locke previously, referred to rights that individuals enjoyed by virtue of their humanity, not as grants or gifts from governments. Indeed, Spencer believed that these natural rights to life and property were antecedent to government and that it was for the purpose of better securing such rights that governments were created. Had governments not been formed, Spencer elaborated, the weak would be vulnerable to threats by the strong, and strong individuals would be at risk for incursions on their property by groups. In response to such threats, individuals joined together to create governments to protect their natural rights to life and property, Spencer concluded.

The origin of government pointed to its purpose. Created to better secure natural rights, government possessed no legitimate purpose beyond this delimited objective. Justice, Spencer (1843/1994d) insisted, was “comprehended” by nothing more than the defense of natural rights. In a later essay (1881), he reaffirmed the definition of justice given in *The Proper Sphere of Government* by referring to justice as the prohibition of actions that “trespass” against another.

Spencer’s understanding of government’s role was logically consistent with his assertion that human existence was subject to the workings of natural laws. Spencer (1843/1994d) taught that human life, like the life of “every animate creature” (p. 48), consisted of the interactions of needs and responses to needs. Living beings had immediate and long-term needs and possessed also “instincts” that impelled them toward the satisfaction of needs. If the environment were altered to thwart the instinctual or organic response, or if it were manipulated to separate the response from its natural outcome, the creature’s existence would be undermined. Thus, whether the being in question happened to be a human being or “the meanest zoophyte,” its natural instincts should not be artificially disrupted. In the case of humans, instinct and intelligence, if consulted, prompted individuals to preserve themselves through multiple “endeavors” (pp. 48–49). In his essay “The Sins of Legislators” (1884/1994e), Spencer explicitly linked this process of instinctual or organic development with evolutionary language that posited “the beneficent working of the survival of the fittest” and cautioned that “aid given to the inferior by the superior” would “enable the inferior to multiply” and produce “mischief” (pp. 131, 128; see also Spencer, 1881, p. 82).

It was nature’s law that individuals who were adept in pursuing their self-interest through self-preserving “endeavors” would be rewarded whereas those who neglected such would suffer. It was not necessary—on the contrary, it was harmful—for governments to attempt to “interfere” with the workings of this natural law. Governments were not, for example, to save people from the consequences of their own poor decisions, nor were governments to direct or guide individual philosophical or

religious beliefs. As Spencer (1843/1994d) put it, government existed to protect life and property but not to regulate trade, provide public education, promote religion, provide assistance to the poor, or maintain systems of public transportation. To do more than protect natural rights was to treat citizens as children. To try to “mould” people—as though people were “dough”—to act in ways that transcended nature’s dynamics, or through one of the “Acts of Parliament” to legislate as though humans could be “twisted” into any shape desired by government actors, was to inappropriately expand the role of government beyond the mere protection of natural rights.

Spencer recognized that some individuals could be said to experience “suffering” under the conditions prevailing under limited government. In his essay “The Coming Slavery” (1884/1994a), he addressed directly the question of whether public assistance—that is, an expansive state authority equipped to implement social welfare programs to help the suffering and the needy—might not be called for in such cases. His conclusion reaffirmed his argument for limited government and his opposition to welfare-oriented public policies. He offered five specific arguments in support of his conclusion. One, suffering—or the desire to avoid it—often proved to be a compelling motive to work hard as an individual, and thus suffering, in and of itself, did not constitute something that merited automatic eradication or prevention. In short, “much suffering is curative” (p. 90), Spencer reasoned. Two, even if one were to decide that suffering constituted an “evil,” evils were not necessarily preventable; therefore, limited government continued to be justifiable even in the presence of suffering. Three, even if some evils did not prove intractable to those who would seek to eliminate them, it was not logical to assume that the mere fact of intractability was the equivalent of establishing a positive obligation on the part of government to institute preventative policies. Four, once governments began to address suffering, a precedent for doing so would be established, and public expectations for continued assistance would be generated; consequently, the transformation of government beyond its original purpose of the preservation of basic natural rights would grow uncontrollably. Fifth, “officialism” would come to characterize the government. *Officialism* was Spencer’s term to describe a highly regulative state, a state—insofar as it was regulatory rather than limited—that had become so large and intrusive as to infringe on the natural rights to life and property, the very rights that government had been created to protect.

In many of his works, Spencer acknowledged that the concept of protection—that is, protecting natural rights—was highly interpretable and subject to a variety of readings. One could argue that Spencer’s neoclassical liberal theory was an exercise in searching for ever more precise and narrow ways to define protection so that only the most minimal level of governmental activity possible would be

allowable. Spencer (1843/1994d) crafted his narrow definition of protection, first, by pointing out that the notion of a “general good” was intellectually flawed. It was not compelling, he asserted, to define the protection of natural rights as the entailment of a general good. The notion “general good,” he continued, was so broad as to be unusable: It could mean whatever the person using the phrase wanted it to mean, and thus, in reality, it had no reliable meaning. It was erroneous, therefore, to say that government, as part of its duty to protect natural rights, had a general obligation to promote a “general good.”

In addition, the proposition that government should promote a general good was mistaken insofar as it rested on erroneous assumptions about human nature and the human capacity to preserve the self. The argument that protection should be interpreted broadly enough to encompass the idea that government should actively promote or produce the “general good” as part of its protection of natural rights rested on the assumption, Spencer (1843/1994d) claimed, of human incomprehension and incompetence as regards self-interest. A stance such as the above accepted the false proposition that good outcomes were unrealizable in the absence of governmental programs designed to directly promote them. Spencer tried to strengthen his case with reference to religious examples. Humans were not so irredeemably one-dimensional, he argued, as to universally reject religion outright in the absence of a state-sponsored church, nor was Christianity so unappealing as to require the laws of the state to force it on the citizenry, nor were unbelievers so innately corrupt as to pose a threat to the natural rights of others by the mere fact of their religious nonbelief and nonaffiliation. Good outcomes—the nonviolation of the natural rights to life and property—could thus ensue without the expansion of state authority into the sphere of religious regulation in the name of serving the “general good.”

Spencer (1843/1994d) also pointed to the example of public health regulations. He argued that advocates for government-sponsored public health protections were assuming citizen incapacity. Such reformers were advocating government involvement in the provision of health care services, Spencer charged, because they saw women and men as too ill-equipped to make rational, discerning choices for themselves. He compared the consumer who would sacrifice quality for lower cost when selecting a pharmacist to one who would do likewise in selecting a watch repairer: In both cases, the individual was said by Spencer to be capable of evaluating risks and rewards and stood in no need of a regulatory state to protect him or her from poor choices. If suboptimal choices were made, the individual in question would have to live with the outcome of having exercised “his own free will” (p. 47), whether making decisions that failed to fix a watch or failed to provide safe, effective drugs.

Finally, Spencer (1843/1994d) critiqued the idea of active government promotion of the “general good” by

concluding that one should construe the violation of natural rights in what he labeled “positive” terms. That is, government’s task was to protect natural rights against direct (positive) violations, not against indirect (negative) effects or by-products. To call indirect effects “infringements” was to fail to understand the meaning of rights *per se*. Spencer used economic examples to establish his distinction. What if, he reasoned, an individual claimed that his or her impoverishment constituted a violation of his or her natural rights; if such a claim possessed credibility, the state—in its duty to defend natural rights—might be understood as having an obligation to enact laws for the provision of economic assistance. However, such claims would lack credibility as long as violation was understood as that which was constituted by a direct act (positive act) rather than as an unintended, indirect act. An individual suffering from poverty, Spencer (1843/1994d) went on to suggest, was simply experiencing the effects of nature’s “self-adjusting principle,” not a positive violation of natural rights (p. 6). The “self-adjusting principle,” Spencer explained, referred to nature’s law whereby productive activity was rewarded and unproductive activity was not. Spencer upheld the benefits of the principle in the case of individuals suffering poverty as a result of their own bad decisions as well as those suffering as a result of bad luck. In neither case did Spencer believe that government should step in and provide assistance to the poor in the name of “general good.”

Indeed, to enact laws for public assistance was to set in motion a series of deleterious effects, Spencer (1884/1994e) warned. From this misconstrued notion of government’s proper role emanated problems ranging from (a) the undermining of incentives both for hard work and for voluntary charity extended to the hard-working-but-down-on-their-luck poor to (b) the confiscation of the wealth generated by the productive members of society. Taxes to fund public assistance programs would fall not only on the affluent, Spencer noted, but also on the productive members of the working class; resentment toward such confiscatory programs would foster antipathy to the poor generally, without distinguishing between those who were poor as a result of misfortune beyond their ability to address and those poor as a result of ignoring nature’s laws regarding individual responsibility for self-preservation. Moreover, individuals who could look to government for public assistance could be misled into thinking they should ignore nature’s “self-adjusting principle” whereby those who care for themselves progress while those who shun this principle do not. Moreover, Spencer (1843/1994d) calculated that “nine cases out of ten” (p. 17) involving poor people, on inspection, would reveal that the poverty in question was merited, either by the actions of the poor themselves or by the actions of their parents. With regard to the latter, it was beneficial, Spencer wrote, for the consequences of parental misdeeds to be felt by children in order to give

the children a compelling incentive not to follow in the missteps of their parents.

Not only was government to refrain from creating antipoverty programs, but government was also to abstain from providing for public education. Spencer's (1843/1994d) neoclassical liberal opposition to public education was based on a variety of arguments. On the simplest level, he maintained that it made no greater sense for a state to provide "intellectual food" (p. 33) than literal food to its citizens, citizens who were responsible for feeding themselves as adults and responsible for feeding their children if they were parents. In addition, he rejected calls for state-sponsored education on the grounds that such education would necessarily, he believed, introduce a "uniform" education on a populace with diverse needs, interests, backgrounds, and capabilities. He also predicted that public teachers would be less responsive to parents and less concerned with achieving high standards than would teachers who did not look to the state as their employers. A state-directed educational system would thus impede national progress. Additionally, however, Spencer raised the question of whether England's prevailing educational theory was sound. If it proved to be deficient, he asserted, to further support that educational theory by providing it state sanction would exacerbate—not rectify—existing problems. What incurred his scorn was the following: an approach to education biased in favor of classical schooling. Specifically, Spencer believed that England's approach to educating its children was oriented toward "dead languages" and old pedagogy. That which Spencer (1843/1994d) termed "dead languages" (p. 38; i.e., classics) were taught at the expense of modern science and modern ideas. Any progress in moving educational methods and curriculum in the direction of teaching modern "real knowledge" would be stalled if the present curriculum were turned into a public education program with state support behind it, Spencer warned. Moreover, as with other hierarchical institutions, a centralized educational system would promote passivity in its students and would also impose tax burdens on the entire population, regardless of agreement or disagreement with its principles and regardless of whether, among those taxed, there might be individuals with no children in the public school system.

Spencer's (1843/1994d) neoclassical liberal vision of minimal government had implications for foreign policy and international relations as well as for domestic policies such as education and economics. In short, Spencer considered war-making powers suspect, and he concluded that only wars to defend against direct attacks on a state's sovereignty were justifiable. Other types of wars were entirely illegitimate for the same reason that government encroachments into domestic domains were unjustified. In both foreign and domestic policy arenas, governments that reached beyond the limited role of defending natural rights were exceeding their appropriate authority. Spencer

believed that a study of the history of warfare revealed that most wars were "evils" (p. 23). He compared war to an intoxicant that had the ability to stupefy the consumer—in this case an entire nation—into thinking that war could bring prosperity and security rather than debt and death. Whatever the rhetoric of politicians, Spencer continued, most wars represented ambitious schemes for "invasions and conquests" (p. 23). To prevent the evil of war, it was necessary to handcuff politicians, to "confine" their powers to the narrowly defined protection of justice discussed above. He urged his native England to set an example for other nations in abjuring war, and he cautioned England not to wait for international consensus before renouncing all wars other than those fought for defensive purposes.

Not surprisingly, given his position on war, Spencer (1843/1994d) was steadfast in maintaining his support for limited government against those who championed colonialism. Indeed, his neoclassical liberalism had a strong anticolonialism element. Spencer argued that states could not be both limited in size and scope and simultaneously committed to the pursuit of colonial expansion. Pointing to the examples of Canada, the United States, and India, Spencer also argued that colonies had imposed costs on England while creating profits only for the "monopolists" and "aristocrats" who invested in them.

For Spencer (1884/1994c), whether in foreign or domestic applications, liberalism was a philosophy that sided with "individual freedom" and against "state-coercion" (p. 66). In his 1884 essay "The New Toryism," Spencer discussed at length his understanding of the word and the philosophy *liberalism*. In this essay, he made the claim that liberalism across the ages had been animated by the goal of reducing "the coercive power of the ruler over the subject" (p. 65). Historically, this goal had guided efforts to achieve judicial independence, the establishment of habeas corpus rights, and the emergence of a Bill of Rights. In each case, liberty was understood in reference to what Spencer termed "negatively coercive" acts; that is, liberals supported a definition of liberty whereby liberty was limited only by the negative prohibition on "aggressing" against another party. In this understanding of liberty, an individual was considered free as long as he or she was not being restrained by government, and it was understood that government would impose no restraint unless the individual initiated an aggressive move against another party. Liberty had not been understood in relation to "positively coercive" acts, that is, liberty had not been conceptualized as something one was required by government to surrender in order for the government to compel the individual to provide positive support for a public good.

Spencer (1884/1994c) argued that liberals in the late 19th century were suffering from doubt over whether the "negative" understanding of liberty should continue to

prevail insofar as the strong monarchical form of government in England had been replaced by one of “parliamentary authority.” He responded forcefully in favor of continuing to uphold the “negative” understanding of liberty. To be specific, he argued that liberalism’s defining element was not the *form of government* (strong monarchy vs. strong parliament) that it sought to limit but was, rather, the commitment to *limiting the power of government*, whatever form government might assume. Liberal vigilance in continuing to contain governmental power within strict limits was no less urgent in a democratic age, Spencer believed, in part because of the role of political parties. Representative government was generally organized along party lines, and parties were often led by “unconscientious” individuals driven by electoral pressures for quick fixes to perceived social ills (e.g., poverty) by attempting to undo nature’s “self-adjusting principle” (Spencer, 1884/1994a, p. 93). Majority opinion was not a legitimate basis for decision making when majorities were calling for expansive government that would replace negative liberty with the “officialism” described above (Spencer, 1884/1994b). The liberalism of the present, Spencer (1884/1994c) maintained, should thus “measure” liberty by the same standard that liberalism had used in monarchical times: Liberty should be maximized, and liberty would only be maximized by the “relative paucity of the restraints it [government] imposes” on the individual (p. 77).

Spencer’s liberalism was critiqued by a number of 19th-century socialist writers. Socialist critics’ arguments tended to challenge Spencer to see that inequalities in economic conditions could be just as oppressive as governmental institutions. Not surprisingly, Spencer regarded such critiques as being as ill-conceived as the classical educational pedagogy to which he, as discussed above, also objected. Just as England’s schools were suffering under the “bias” of the “dead languages” of antiquity, so did socialist calls for freeing modern individuals from what socialists described as economic tyranny unconsciously draw on ancient ideas, Spencer wrote. That which Spencer considered to be a dead ancient idea was classical Greek political theory upholding the polis as a source of the individual’s very identity. The Greek ideal, as Spencer (1885/1994a) saw it, made the “individual . . . a slave to the community” (p. 103). Spencer saw socialism as an effort to implement a resurrected Greek polis ideal. Socialism was nothing short of “slavery” and was a throwback to antiquity’s celebration of the idea that the polis, or society, should be dominant in relation to the self.

Spencer’s (1884/1994e) neoclassical liberalism praised *market forces*—what he called “the spontaneous activities of citizens, separate or grouped” (p. 125)—for not only constituting the conditions under which individual liberty could be maximized but for also generating economic progress. Markets unleashed “men’s desires,” and individuals pursued those desires through, at their discretion, “private activities” as well as “spontaneous cooperations” to

produce unprecedented agricultural, manufacturing, and technological progress. Spencer pointed to developments ranging from the invention of the telephone to modern science as due “not [to] the State” but to “the aggregate results of men’s desires” (p. 125). Indeed, states, Spencer believed, were more likely as not to have obstructed progress by regulating human activity whenever states entered into economic affairs.

In sum, Spencer’s neoclassical liberalism championed liberty, not a particular class, partisan position, or governmental structure. When “aristocrats” and “monopolists” were the opponents of liberty (as in the case of colonialism), he criticized these elite groups. When kings attacked liberty, Spencer criticized monarchy. However, when the opponents of liberty were reformers advocating for working-class benefits or parliamentarians using a reform-oriented rhetoric, Spencer criticized reformers and parliamentarians.

William Graham Sumner

One of Yale’s most popular 19th-century professors, U.S. sociologist William Graham Sumner, is, like Spencer, generally regarded as both a neoclassical liberal and a social Darwinist (Hofstadter, 1955). Sumner’s neoclassical liberal political theory was given explicit expression in his 1883 publication titled *What Social Classes Owe to Each Other*. In this work, Sumner analyzed government not only in terms of its role but also in reference to what he regarded as its essence. The role of government, he contended, was nothing other than the protection of “the property of men and the honor of women” (Sumner, 1883/2007, p. 58). That is, governments were “contractual” and “rationalistic” responses to individual needs. Governments were instituted to improve the security of life and property, and, as such, governments existed for reducing threats against persons and property. The government, in keeping with neoclassical liberalism’s logic, was to serve this minimalist purpose in a very strict fashion and was, in particular, to avoid “sentiment” toward those less fortunate or those who might need coaxing out of their present choices. To Sumner, any government intrusion into “sumptuary” or “moral” areas for “legislation” (e.g., prohibitions on alcohol consumption or taxes on items in order to curb the consumption of those items) was unjustified. Legitimate governments were neither “paternal” (providing economic or moral safekeeping to citizens) nor “protectionist” (protecting individuals from the effects of competition or from the effects of their own poor choices). For Sumner, anyone who would argue convincingly in favor of expanding the role of government beyond its minimalist scope would have to prove that it would be “more advantageous” to those who would “bear the weight” of the expansion (e.g., those taxed to support government programs) than would

“complete noninterference by the State” (p. 82). Not having encountered any such proof, Sumner concluded that, in maintaining a strict contractualist, minimalist government, society was best promoting “the self-reliance and dignity of a free man” (p. 21).

Moreover, for Sumner (1883/2007), states or governments were not an embodiment of some larger-than-life repository of wisdom or expertise that could be called on to mandate superior planning and oversight for societal interactions. States were “in practice” nothing more than collections of “obscure clerks” who, in democratic republics such as the United States, were put in their positions “in a very haphazard way by the majority” (p. 12). For example, states did not have the wisdom to “make men happy” but had only the purpose of providing a secure setting in which individuals made for themselves lives that were happy or not (p. 25). Like Spencer, Sumner applied his neoclassical liberal commitment to limited government in the area of foreign as well as domestic policy, and like Spencer, Sumner criticized “imperialist” and expansionary military ventures as inconsistent with liberalism (Sumner, 1899).

Sumner’s advocacy of minimal government was based on his understanding of nature, rights, individualism, and science. With respect to nature, Sumner’s neoclassical liberalism rested on four observations. First, nature was characterized by a scarcity of goods, for which individuals were in competitive struggle. Sumner (1883/2007) conceptualized nature as having no “pity” for the unfortunate. He described human existence as “naturally” characterized by “hardships” and “struggles,” struggles in which positive outcomes are guaranteed to no one (p. 17). In illustrating his thesis, Sumner wrote of two hypothetical “neighbors” and set up an imagined scenario in which one neighbor turned out to be more successful than the other; Sumner then remarked that the less successful neighbor had no legitimate basis for ill will toward his more successful peer because the former could not rightfully “blame” another for what (unequal outcomes in a struggle for scarce goods) was only natural.

Second, Sumner’s belief in nature and natural struggle did not lead him to deny a role to cultural and historical influences on human life. As cultures evolved, new ideas accompanied new ways of organizing social life. For example, Sumner (1883/2007) pointed to modernity’s repudiation of feudal notions of inherited social status. A careful student of history, he argued, could see that the neoclassical liberal theory of limited government was itself a modern invention that represented a rupture with medieval ideas that had supported the expansive powers of elites over subjected populations. “The notion of a free state,” he observed, “is entirely modern” (p. 24). That is, Sumner’s emphasis on the inescapability of nature’s forces was put forward in a context that also recognized societal development (e.g., the movement from feudalism to capitalism) as having

played a part in shaping societal ideas and individual choices. That he attributed an influence on human interactions to culture—not to nature alone—has led some scholars to suggest that Sumner’s liberalism owed less of an intellectual debt to 19th-century evolutionary science and social Darwinism than has often been assumed (Smith, 1979).

Third, nature imposed penalties for poor choices, and poor choices were defined as those that inhibited long-term security, Sumner taught. Any individual, Sumner asserted, was capable of making poor choices. Indeed, Sumner stated explicitly that “human nature” was replete with “vices and passions,” and he named among them such vices as “cupidity, lust, vindictiveness, ambition, and vanity” (Sumner, 1883/2007, p. 23). Any individual in any age, he elaborated, could manifest any such vice. However, some individuals followed “reason and conscience” and thus overcame temptations to indulge vices; such individuals chose to practice “self-denial” and “self-control.” Sumner referred to those who exhibited such salutary behaviors and attitudes as the “elites.” Those who indulged vices would, if a natural competitive struggle ensued, be more likely to succumb than those who practiced self-control. Sumner offered the example of a “drunkard” and the example of a gambler. If left unobstructed, nature would “remove” those who rendered themselves vulnerable through their indulgence of vices. Comments such as these contributed to Sumner’s reputation as a theorist who was no less a social Darwinist than a classical liberal.

Fourth, Sumner (1883/2007) argued that natural outcomes could be skewed in terms of their directions but could not be obviated altogether. Sumner contended that one could “divert” a consequence, but one could not preclude it. He pointed again to the example of the drunkard. If a drunkard were to be spared the natural consequences of his lack of self-control, the penalty for drinking would, in this case, have simply been shifted to the party that made possible the sparing. For example, Sumner elaborated, were a police officer to intervene in a drunkard’s life to rescue him from the streets, any person earning a wage or salary that was taxed to pay for the police officer’s intervention was being made to endure (insofar as taxes were compulsory) the costs of the drunkard’s vice. In this instance, the drunkard was getting away with vice rather than facing the consequences of his own poor choices, and the productive tax-paying citizenry was being penalized.

Sumner (1883/2007) referred to the latter as the *Forgotten Man* or, occasionally, as the *Forgotten Woman*. Sumner characterized the Forgotten Man or Woman as productive, so productive that he or she served as an easy target for taxes needed by social reformers acting on “sentiment” for the drunkard (or other needy individuals or groups). Sumner described the Forgotten Man as forgotten in the sense that—except as a tax base—he was invisible to reformers who, out of concern for the drunkard, the gambler, the poor, or others, needed a tax source to fund

reform-oriented expansive government assistance programs. The Forgotten Woman, Sumner explained, might be a seamstress whose supplies were more expensive than they would otherwise be as a result of reformers' efforts to "protect" an industry and its workers from competition. Although such reformers might appear big hearted and generous when one looked at their avowed concern for the vice-oriented or unfortunate people they wished to use the government to assist, such reformers were somehow, Sumner argued, unable to see the injury inflicted on the Forgotten Man by confiscatory taxes levied against his earnings and savings or the harm visited on the Forgotten Woman by the regulatory state. It would have been inconvenient for reformers to acknowledge the Forgotten Man's or Woman's existence, for doing so would have made it clear that reformers were not eradicating problems but were, rather, simply shifting problems and (tax or protectionist) penalties onto those who did nothing to deserve them.

In addition, private acts of charity toward a person who exhibited "shiftless" or "inefficient" attitudes or behaviors similarly skewed the direction of natural penalties, Sumner (1883/2007) believed. Just as government programs to help the "drunkard" and those like him were unjustifiable violations of the concept of minimal government and resulted in injury to the Forgotten Man and Woman, so were private disbursements of charity expenditures that, in being given to the vice indulgers, were distributed away from more productive areas of investment. Specifically, an amount given in charity to a drunkard was an amount not invested in a business to employ those—such as the Forgotten Man or Woman—who were hard working and who practiced self-control; nor was this amount contributed toward the savings of one's own family so that the family might be better protected against the possibility of unforeseen needs.

In contrast to Spencer (and Locke) before him, Sumner (1883/2007) based his neoclassical liberal arguments in favor of minimal government, not on an appeal to the self-evident nature of natural rights, but, rather, on a critique of the concept of natural rights. Sumner argued that to say that individuals are born with natural rights tends to imply (a) an entitlement to that which is promised by the right and, further, to imply that there is (b) an obligation held by someone to recognize the right. Sumner rejected both ways of thinking about rights. If when discussing rights one were referring to nothing more than "chances," he contended, he did accept the concept of rights. If, in saying that all individuals have the right to property, one is merely claiming nothing more than that all individuals have a chance to struggle and compete for property, Sumner could then agree with a rights-based argument. These rights or chances should be "equal," Sumner contended, and from this equality of opportunity, "the merits of individuals" (p. 88) would determine outcomes. His neoclassical liberalism

emphasized, however, the misuses to which a language of rights could be put, insofar as individuals who thought of themselves as having natural rights might automatically assume that either other individuals or the government had enforceable obligations to ensure the realization of those rights.

Sumner's (1883/2007) individualism—that is, the belief that individual liberty should be maximized as long as it did not entail aggression against another—was put forward in a context that recognized (a) social duty and (b) group life as elements of human existence. With respect to an individual's duty to society, Sumner maintained that each individual had a duty to care for himself and, if he had a family, to care for his family. In caring for himself and his family, the individual was meeting his social obligations in the sense that he was allowing neither himself nor his family to become a burden on others. Thus, the individual's likely existence within a family unit was folded into Sumner's very notion of individualism. Moreover, Sumner argued that individuals sometimes found it rational to pursue their individualism as members of nonfamilial groups, or "associations." For example, he asserted that members of the working class might, under various conditions, find it rational to combine in "associations" or unions to compete for their respective individual economic interests, even though he believed that U.S. workers were too "independent" to provide sustained support for unions. Monopolistic industries that created "duress" for consumers or employees by reducing liberty through the curtailment of competition were opposed by Sumner (Curtis, 1969).

Sumner (1883/2007) presented his neoclassical liberalism as an approach informed by "science," a science that dispassionately assessed natural and social forces. Sumner's understanding of science was such that he regarded it as "impersonal" in its observations and conclusions. For example, just as an impersonal analysis of nature's forces demonstrated the existence of competitive struggle and the enhanced survival chances for those practicing "self-control," so did a scientific study of society demonstrate the benefits of "laissez faire." Laissez-faire doctrine could be summed up, Sumner asserted, as "mind your own business" (p. 67) and do not try to use government to impose regulations on another person's life. Thus, he explained, the task of social scientists was not to admonish others in the language of "ought" and sentimentality but to communicate, instead, in the language of cause and effect, analysis and explanation.

Yet Sumner (1883/2007) realized that individuals were not reducible to rational calculations and that individuals were often connected to one another by bonds of "affection" and "sympathy." He knew, for example, that individuals might be moved by "sympathy" for another person's situation and might, out of compassion, act to alleviate someone's distress, even though such actions might make no sense from the viewpoint of "impersonal" scientific

assessments. Viewing someone from the standpoint of sympathy yielded information that lacked the rigor of science, but such a view offered insights nonetheless, Sumner realized. Sympathy, he wrote, could point individuals in the direction of acknowledging that even “the best of us act foolishly” at times. There were circumstances in which each might gain if “we share each others’ burdens” (p. 86).

Sumner (1883/2007) did not regard neoclassical liberalism’s objective as one of universally condemning actions undertaken from sympathy, but, rather, he saw liberalism’s logic as asserting that such expressions were best conceptualized as “two party” interactions. That is, the terms and conditions of such interactions should be at the discretion of the two parties immediately engaged (e.g., between the one offering help and the one receiving help; between the one giving affection and the one receiving affection) and not mandated by a third party (e.g., the state). In short, although an individual who never felt sympathy for another would be a lamentable human being, Sumner concluded, a government that forced citizens to act on the basis of sympathy would destroy individual liberty.

Conclusion

In their political writings, both Spencer and Sumner drew on 17th-century philosophical ideas (i.e., Lockean support for a limited government and for maximized individual liberty), incorporated 19th-century scientific currents (i.e., evolutionary scientific perspectives on nature as an arena of struggle), and commented on various social and political issues of the day (e.g., outdated pedagogies in the British school system in the case of Spencer and reformist demands for tax-supported social assistance programs in the case of Sumner). As theorists who assessed and responded to issues of their times, these men penned broad theoretical generalizations immediately accessible to readers across the centuries, as well as historically specific examples that may seem archaic to contemporary readers. Their theories could be bold as well as nuanced, with the former indicated perhaps by no example more telling than Spencer’s phrase “survival of the fittest” and the latter exemplified by Sumner’s acknowledgment of the utility of workers’ associations in pursuing individual self-interest and the role of sympathy in everyday human interactions. As “new” classical liberals, Spencer and Sumner shared Locke’s belief that individuals had sufficient incentives to refrain from harm to others in the absence of an expansive “big” government that existed to regulate individual choices and behaviors. Whereas Locke had pointed to the role of reason as an internal source of self-restraint, Spencer and Sumner emphasized nature’s harsh incentives, especially as reinforced by a limited government that would confine its role to the protection of private property

and life from direct attacks and would not attempt to regulate out of existence—on the basis of some presumed government wisdom—the effects of what Spencer called nature’s “self-adjusting principle” and what Sumner called “laissez faire.”

While Spencer and Sumner were dominant voices within 19th-century neoclassical liberalism, the core concerns of this school of liberalism are by no means confined to these two writers. The neoclassical liberal emphasis on limiting the power of government is found also in the 19th-century utilitarian liberalism of John Stuart Mill (1806–1873), in the arguments of contemporary neoclassical liberal theorist Robert Nozick, and in the platforms of the Libertarian Party. In ways reminiscent of Spencer and Sumner, Mill upheld freedom of expression against the claims of governmental censors, Nozick champions a Lockean umpire state, and Libertarians advocate for the rights of individuals to claim their own sexual identity (e.g., gay rights) and to make their own reproductive choices (e.g., abortion rights). Over the centuries, the specific areas of public policy capturing the attention of neoclassical liberals have changed, but the vision of Spencer and Sumner endures. In short, although 21st-century neoclassical liberals have generally tried to soften or expunge the social-Darwinist implications of neoclassical liberal theory, contemporary advocates of the legacy of Spencer and Sumner continue to advocate on behalf of the benefits of minimal government and maximum liberty for rational self-interested individuals.

References and Further Readings

- Ball, T., & Dagger, R. (2009). *Political ideologies and the democratic ideal* (7th ed.). New York: Pearson Longman.
- Curtis, B. (1969). William Graham Sumner “On the Concentration of Wealth.” *Journal of American History*, 55, 823–832.
- Hofstadter, R. (1955). *Social Darwinism in American thought*. Boston: Beacon Press.
- Kramnick, I. (1998). Equal opportunity and the “race of life.” In N. Love (Ed.), *Dogmas and dreams: A reader in modern political ideologies* (2nd ed., pp. 99–112). Chatham, NJ: Chatham House.
- Miller, W. L. (1975). Herbert Spencer’s optimum development path. *British Journal of Sociology*, 26, 477–485.
- Offer, J. (1994). Introduction. In J. Offer (Ed.), *Spencer: Political writings* (pp. vii–xxix). New York: Cambridge University Press.
- Pickins, D. K. (1981). Western expansion and the end of American exceptionalism: Sumner, Turner, and Webb. *Western Historical Quarterly*, 12, 409–418.
- Smith, N. E. (1979). William Graham Sumner as an anti social Darwinist. *Pacific Sociological Review*, 22, 332–347.
- Spencer, H. (1881). Replies to criticisms on the data of ethics. *Mind*, 6, 82–98.
- Spencer, H. (1994a). The coming slavery. In J. Offer (Ed.), *Spencer: Political writings* (pp. 80–105). New York: Cambridge University Press. (Original work published 1884)

- Spencer, H. (1994b). The great political superstition. In J. Offer (Ed.), *Spencer: Political writings* (pp. 140–169). New York: Cambridge University Press. (Original work published 1884)
- Spencer, H. (1994c). The new Toryism. In J. Offer (Ed.), *Spencer: Political writings* (pp. 63–79). New York: Cambridge University Press. (Original work published 1884)
- Spencer, H. (1994d). The proper sphere of government. In J. Offer (Ed.), *Spencer: Political writings* (pp. 1–57). New York: Cambridge University Press. (Original work published 1843)
- Spencer, H. (1994e). The sins of legislators. In J. Offer (Ed.), *Spencer: Political writings* (pp. 106–139). New York: Cambridge University Press. (Original work published 1884)
- Simon, W. M. (1960). Herbert Spencer and the “social organism.” *Journal of the History of Ideas*, 21, 294–299.
- Stark, W. (1961). Herbert Spencer’s three sociologies. *American Sociological Review*, 26, 515–521.
- Sumner, W. G. (1899). The conquest of the United States by Spain. *Yale Law Journal*, 8, 168–193.
- Sumner, W. G. (2007). *What social classes owe to each other*. Charleston, SC: BiblioBazaar. (Original work published 1883)
- Werhane, P. H. (2000). Business ethics and the origins of contemporary capitalism: Economics and ethics in the work of Adam Smith and Herbert Spencer. *Journal of Business Ethics*, 24, 185–198.
- Wolfe, A. (2009). *The future of liberalism*. New York: Alfred A. Knopf.

71

MODERN DEMOCRATIC THOUGHT

JEFFREY D. HILMER

University of British Columbia

Democracy is a venerable idea. Invented in ancient Greece more than 2,500 years ago, the idea of democracy continues to exercise a powerful influence over the modern mind. The theories and practices potentially included in the expansive category “modern democratic thought” are numerous. In this short chapter, no attempt is made to be comprehensive. Rather, the field of inquiry is strategically narrowed to present a concise overview of modern democratic thought. The assumption that guides the organization of this chapter is that the reader is interested in academic political science and possibly curious about pursuing that interest further. Therefore, in this chapter, the ostensibly overwhelming topic of modern democratic thought will be tackled first by situating the concept of democracy within the context of American political science. Once our subject has been so contextualized, the functional distinction between democratic thought and democratic theory will be introduced. Next, the category of democratic theory will be subdivided into five categories, each representing one of the major theories of democracy extant today. The often fuzzy relationship between democratic theory and practice will be touched on and followed by some pertinent avenues for future research.

Political Science and Modern Democratic Thought

Systematic inquiry into democracy is as old as Aristotle’s investigation into ancient Greek constitutions and as recent as a research paper presented at the last meeting of the American Political Science Association (APSA). Since their Associations’ inception in 1903, American political scientists have been keenly interested in the subject of democracy. As chronicled by John G. Gunnell (2004), APSA has always consciously and unconsciously understood its mission in terms of democracy, whether it was to describe how democracy functioned in the past or how it functions in the modern present, to predict how it may function in the future, or to prescribe how it ought to function. The development of democratic theory as an academic subfield of political science coevolved with that discipline. From the time of its founding, the theory of the state dominated the literature of what we would now call modern democratic thought. In the early 1930s, the theory of pluralism came into vogue. This theory was eventually supplanted by the theory of liberalism. In the 1950s, pluralism was (re)articulated in the context of internal methodological and ideological debates within APSA.

The debate among American political scientists is usually described as taking place between those who sought to transform political science into an empirical enterprise approximating the rigor of the natural sciences and those who saw political science as a more subjectively normative enterprise. Although this dichotomy is somewhat misleading, it adequately describes the disciplinary milieu in which American political scientists during the 1950s, 1960s, and 1970s saw themselves. Subsequent theories of democracy developed in the 1980s onward have in one way or another been a reaction to the hegemony of pluralist theory. The *theory of pluralism*, developed chiefly by Robert A. Dahl (1956, 1961), reflects that debate in that it was understood to be an *empirical* theory of democracy, but also, somewhat paradoxically, to be employed for the very *normative* end of understanding and thus promoting American-style democracy at home and abroad. American political scientists have fundamentally shaped how scholars and nonscholars alike will think about democracy in the 21st century.

Varieties of Modern Democratic Thought

The distinction between democratic *thought* and democratic *theory* is admittedly a slippery one. It is, however, a distinction that can help us understand how political scientists approach the subject of democracy.

Democratic Thought

Democratic *thought* may be defined as a general category that encompasses less systematized political thinking. According to this definition, literature, historical works, political pamphlets, and sundry examples of loosely organized thoughts on popular participation in government may be described as democratic thought. Examples of democratic thought abound. The political thought of James Madison in the *Federalist* (Hamilton, Jay, & Madison, 1966) is an attempt to describe how the proposed U.S. Constitution of 1787 would be structured, explain how that Constitution would address the problem of faction, and predict how that Constitution would balance factional interest in a way that would enable the republic to endure. This is, no doubt, systematic political thinking of a kind. Yet even though democratic thought is often systematic, it does not usually measure up to the meticulous standards of modern political science. As every political science and U.S. history undergraduate knows, the *Federalist* is a collection of articles originally published in New York newspapers in an overtly political attempt to sway popular opinion in defense of the proposed Constitution. Modern political science demands empirical and analytical rigor and objective analysis. Therefore, while *political thought* is often seen to be connected to a very real political and therefore often ideological context, *political theory*—often with the prefix *empirical* or *positive* or *analytic*—is

considered to be *mainly* free from ideological biases and more akin to pure scientific theory.

Democratic Theory

Democratic *theory* may then be defined as a more systematic attempt to describe or predict—often both—the behavior of a political phenomenon. Yet given the tension among political scientists as to their discipline's ultimate objective, democratic theories are more appropriately conceived of as existing on a continuum between two poles: on one end *empirical* and on the other *normative*. Few political scientists would claim that a political theory can ever be *purely* empirical. Rather, the objective is to make it as value-free as possible. Still, democratic theory is generally more rigorously systematized than is democratic political thought. Predictably, then, democratic thought often serves as the inspiration for democratic theories. This is definitely the case with Dahl's (1956) conversion of Madison's political thought into pluralist theory.

Most contemporary democratic theorizing takes place within the academy and is the product of professional political scientists. A review of the literature on contemporary democratic theory will promptly reveal that five theories of democracy dominate the literature. While there are numerous theories of democracy, and much overlap among them, the five principal theories existing today are pluralist, participatory, liberal minimalist, deliberative, and agonistic. *Every student who is intent on becoming conversant in modern democratic thought must understand the ways in which these five theories differ.* Most of the theoretical debate among political theorists on topics of democratic theory involves one or more of these theories. Moreover, much of the debate among political scientists in one way or another involves these theories of democracy.

Pluralist Theory

In *A Preface to Democratic Theory*, Dahl (1956) formulates the democratic theory of *polyarchy*. Polyarchy is Dahl's version of pluralism. Pluralist theory in American political thought and political science was not Dahl's invention but rather has its roots in the political thought of James Madison. Building on Madison's ideas, Dahl describes *polyarchy* as government by various groups in the context of a competitive political system in which each group is endeavoring to achieve its narrow objectives. The elementary tenets of polyarchy are basic citizen rights, including the right to vote (including self-government, majority rule, and political equality), freedom of expression and organization, a system of institutional checks and balances, a competitive electoral system with at least two political parties, and the right and ability of citizens to control the political agenda. Dahl subsequently illustrates how polyarchy functions in the American context in his classic study *Who Governs? Democracy and Power in an American City* (1961).

What is significant about Dahl's work is his attempt to develop an empirical style of democratic theorizing consistent with the tenets of the positivistic behavioral approach popular at the time. Dahl assessed the state of democratic theory in the 1950s to be insufficient because there was no consensus as to the objective of democratic theory. Was the objective to describe or predict or shape political behavior? Dahl clearly stated his intention in *A Preface to Democratic Theory* (1956) to critically evaluate the two strains of democratic theory evident in American political history, the *Madisonian* and the *populist*, and, highlighting the weaknesses of those descriptive theories, develop an empirical theory of democracy. Yet evident in much of Dahl's *Preface* is a mixing of objective analysis and a normative defense of polyarchy. Dahl's definitional characteristics of polyarchy reflect this. That Dahl identifies voting as a basic right of any authentic polyarchy is not necessarily unusual. But his recommendation for *direct* citizen control of the political agenda is somewhat inconsistent with pluralist theory; it is inconsistent because pluralists presumed that the political system (i.e., the rules of the political process) allowed citizens *indirect* control of the political agenda through their representatives, who, the pluralists assumed, would always act in their constituents' best interests. Dahl is attentive to the plight of any minority—racial, economic, and so forth—that is shut out from the political process, and he seems to tacitly endorse non-“legitimate” participation in order to achieve access to that process. Dahl attempts to have it both ways: to be an objective political scientist and a partisan of the polyarchic American political system. This combination produces a pronounced tension in his work during the behavioral period. Given the academic context, critics were apt to attack this contradictory character of Dahl's *Preface*.

During the 1960s and 1970s, the theory of polyarchy was thoroughly and effectively criticized. By the late 1970s and early 1980s, both Dahl and his collaborator Charles E. Lindblom had modified their theories. Lindblom (1977) argued that a key axiom of polyarchy was erroneous. Taking cues from neo-Marxist critics, Lindblom argued that the pluralist theorists of democracy were naive in assuming that powerful groups, especially big business, did not exert disproportionate influence over the political process. The “privileged position of business” in the polyarchical system enables corporations to disproportionately influence the political decision-making process on matters that often have profound consequences for the entire society. Correspondingly, Dahl (1986, 1989) acknowledged defects in the early theory of polyarchy. The dilemma of polyarchy is that it is a theory that recognizes the right of individuals and their organization to be independent and autonomous, but it also is aware that they may take advantage of that opportunity to serve their private interests, often at the expense of the public good. It seems clear that Dahl sought to address a key defect in the theory of polyarchy, namely, the assumption that all individuals and groups were roughly politically equal and

could not subvert the public good without expecting an equally powerful countervailing response from other groups. It is increasingly obvious to Dahl that the individuals and groups that most needed to be controlled in defense of political equality and democracies were corporate officials and corporations. Dahl (1986) argued in *A Preface to Economic Democracy* that control over corporate enterprise in the United States ought to be in the hands of ordinary citizens. This would ensure their political and economic independence, which are both prerequisites for political and social equality and, hence, democracy.

As a testament to its enduring allure, the pluralist theory of democracy has been modified once again. Advocates for the “new pluralism”—chiefly William E. Connolly (Campbell & Schoolman, 2008) and Chantal Mouffe (Laclau & Mouffe, 1985; Mouffe, 2000, 2005)—argue that although various aspects of the “old pluralism” are valuable, a new conceptualization of the theory that modifies it in light of postmodernism, poststructuralism, critical theory, and feminist theory is now needed (see the account of Mouffe's agonistic pluralism below).

Participatory Theory

The theory of *participatory democracy* has roots in the direct democracy of ancient Greece. Unlike proponents of the pluralist theory of democracy, who conceived of citizenship and political participation in terms of voting, interest group membership, and other conventional modes, participatory democratic theorists envisioned political participation in a much more expansive sense. Carole Pateman's significant book published in 1971 marked the first comprehensive articulation of the theory of participatory democracy. Building on the political thought of J.-J. Rousseau, J. S. Mill, and G. D. H. Cole, Pateman contrasts participatory theory with pluralist theory and argues that active participation enables citizens to self-develop their defining rational capacities, as well as engendering positive psychological benefits, including feelings of political efficacy. Additionally, Pateman argues that a participatory society requires that the scope of the term *political* is extended to cover spheres outside the national government.

Accordingly, Pateman (1971) describes the system of workers' self-management in Yugoslavia. Pateman intimates the potential social, political, and economic benefits of participatory democracy by offering this exemplar of a political system in which the industrial sector of society has been relatively democratized and citizens have been educated and empowered by the democratic participation they engage in at the workplace. The workplace, as Pateman argued, was thought to be a sphere in which existing undemocratic power structures could be supplanted by worker ownership and democratic decision making. Participatory democratic theorists argued that democratic participation conceived of in a broader way would produce myriad benefits unrealizable by conventional pluralist

modes of democratic participation. These benefits included the following:

- Psychological benefits: Human beings would be able to realize their full potential.
- Political benefits: Citizens would experience a kind of empowerment that would enable them to break out of the apathy that seemed endemic to many political scientists in the 1950s.
- Social benefits: The “private sphere” of society, namely, home and family life, which was considered outside the political sphere, would be democratized.
- Economic benefits: The dangers that inequalities especially economic inequalities pose for democracy are acknowledged and addressed via workers’ direct control over their productive activities.

C. B. Macpherson (1977) further articulated the theory of participatory democracy by proposing a participatory parliamentary or congressional structure. Similar to Pateman, Macpherson argues that the goal of participatory democracy is to enable members of the society to self-develop as human beings and as citizens while working toward a just society. Whereas Pateman made a case for participation in the workplace, Macpherson seeks to democratize already existing traditional political institutions. He envisions a hybrid of direct and representative democracy organized as a pyramidal system, with direct democracy at the foundation and *delegate democracy* at each level above. Although Macpherson is far from convinced that such a system is without its problems, he is nevertheless optimistic because the base of the pyramid rests on a firm foundation of direct democracy in the neighborhood and workplace. Thus, despite the inevitability that the further one travels up the pyramid, the less direct democratic participation, the system remains directly controlled by the citizens who are actively participating at its base. Delegates would be authorized to act in the peoples’ stead at higher levels of the pyramid, but they would be subject to recall if they did not act according to the people’s will. Political parties would likewise be subject to increased control by citizens.

Benjamin Barber’s (1984) *Strong Democracy* remains the most comprehensive statement of participatory democratic theory. On a philosophical level, Barber argues against the need for any transcendental foundations for participatory democracy. On a practical level, he develops a comprehensive theoretical model of a participatory polity. Barber reenvisioned a participatory democracy to encompass all sectors and modes of political society. His vision is similar to, if more nuanced than, both Pateman’s and Macpherson’s. The theme of expansive participation is expounded on by Carole Gould (1988), who argues that democratic decision making can and should transcend the political sphere to include economic and social life, too. Nonetheless, although books on participatory democracy

were being published in the early 1980s, by this time political scientists were losing interest. Increasingly, the *participation* in participatory democracy was understood in a more limited, deliberative sense. Therefore, today the *deliberative theory of democracy* is regularly conflated with participatory democratic theory. As we shall see momentarily, deliberative democracy does involve citizen participation, but it is a mode of participation that is less comprehensive and largely restricted to traditional political sectors of a society.

In spite of attempts at revitalization, the theory of participatory democracy has been superseded by liberal minimalist, deliberative, and agonistic theories of democracy. Although participatory democracy has not completely vanished from the literature of political science, it is ever more challenging to locate democratic theorists who actively further it by reiterating its normative principles, theorize participatory institutions, and search for evidence in support of its feasibility. According to Boaventura de Sousa Santos (2005), in the past decade, the practice of participatory democracy has increasingly become manifest around the globe. This is especially the case in South America. As described by Leonardo Avritzer (2009), the proliferation of participatory institutions in Brazil intimates that participatory democracy is viable. Despite the dearth of interest among American political scientists, it is intriguing to speculate about what they might learn by studying contemporary international manifestations of participatory democracy. Given the participatory trends in Brazil, a renewal of participatory democratic theory may be on the horizon.

Liberal Minimalist Theory

The appropriate text to begin is with Kenneth J. Arrow’s *Social Choice and Individual Values* (1962). Arrow’s is the seminal work of *social choice theory*. Although he does not address political issues directly, he identifies disturbing issues that trouble democratic theorists. Arrow’s famous *impossibility theorem*, or *Arrow’s paradox*, addresses a fundamental problem encountered by any political system that requires three or more choosers, or voters, to register their preference from a list of three or more alternatives in the formulation of public policy. Arrow presents a formal proof logically demonstrating the impossibility of creating any effective aggregation device. At the core of that proof is the second issue of *cyclical majorities preference pattern*. If we make the reasonable assumptions that voters have divergent preferences and that the issue that receives the majority of votes will win, the theory of cyclical voting reveals that the order in which issues are voted on will determine which issue wins. No matter what aggregation device is used, and no matter how fair the system of aggregating individual preferences is, the procedure will always yield an irrational result. Therefore, any attempt to amalgamate the individual preferences of voters into rational social policy is bound to fail.

The preeminent exemplar of the influence of Arrow's thought on democratic theorizing is William H. Riker's *Liberalism Against Populism* (1982). Riker is an influential proponent of rational choice theory in political science and a fervent critic of direct democracy. Understandably, then, his analysis of democracy originates from that specific methodological angle of vision. Riker evaluates the implications of social choice theory generally—and Arrow's paradox specifically—for democratic theory and practice. In Riker's view, democracy refers to both an ideal (participation, liberty, and equality) and to a method or procedure for aggregating individual preferences into collective social polity (voting). It has been assumed that the democratic ideals of participation, liberty, and equality could be achieved via the expression of the general will or common good through voting. But is this practically possible? Riker asserts that this important question has never been effectively addressed, but he also thinks that this vexing question can be answered by applying the analytic theory of social choice to democracy.

Riker (1982) begins his analysis with an assumption consistent with the rational choice theory principle of *methodological individualism*. According to Riker, citizens expressing their preferences at the ballot box are the necessary if not sufficient condition for democracy. Riker is quick to point out the irony. Voting is necessary for democracy, but unless it is simple majority voting between two alternatives—which it rarely if ever is—democracy becomes problematic (Arrow's paradox). Riker expands on Arrow's point by emphasizing the problems of strategic voting and agenda setting. *Strategic voting* occurs when a voter endeavors to use his or her vote in such a way as to render the final outcome of an election contrary to what it would have been if all voters had cast their ballots honestly. Riker uses the example of a plurality system in which a voter who favors a third-party candidate strategically votes for his or her second choice with the intent of defeating the major party candidate he or she opposes. Riker argues that this kind of strategic voting is a universal and “ineradicable” problem. *Agenda setting* is an equally vexing problem. Power to set the agenda is often exercised by party elites and occasionally by ordinary citizens. One example of agenda setting is the selection of candidates or issues that are presented to the voters. Often voters have little say in this preelection selection process, which—as Arrow and Riker note—is necessary for voting to function and yet also highly undemocratic. Because strategic voting and agenda setting are next to impossible to eliminate, and because they render voting results inaccurate and “manipulated amalgamations” rather than “true amalgamations” (Riker, 1982, p. 238) of the voters' will, voting in most democratic political systems is inherently unfair and violates the principles of liberty and equality.

If we assume that voting is the “central act of democracy” (Riker, 1982, p. 5), as Riker does, we understand why he distinguishes between “liberal” and “populist”

interpretations of voting. Voting is a protective device employed by citizens to protect their rights from various forms of tyrannical and antidemocratic encroachment. The liberal minimalist does not assume that there is a “general will” or “common good” that is expressed by the people at the ballot box. Rather, democracy is a method by which citizens may exercise control over their leaders, thus preventing tyranny and preserving liberty and political equality. Furthermore, the liberal interpretation of voting is also consistent with the rational choice approach in that the means are considered rational and normatively valuable, but no judgment is made about the ends. The features of the liberal theory of democracy and voting are brought into relief when juxtaposed with its populist counterpart. The populist argues that participation in politics means participation in formulating and enacting legislation. Far from a negative function, the populist—who in Riker's view naively assumes that the majority will not tyrannize—argues that participation is necessary for liberty. Indeed, liberty is the product of that participation, rather than, in the liberal view, only a means by which to protect it.

Riker's (1982) solution to the problem of democracy and Arrow's paradox is the rejection of populism and the embrace of liberalism. Why must populism be rejected? As we have already noted Riker's theory of social choice suggests that it is impossible for the popular will to be amalgamated into a clear policy direction by voting. Instead, in a sleight of hand, Riker replaces the practice of majority voting with the idea of the popular will. There is, then, no possibility of majority tyranny because there is no majority to tyrannize the minority. This explains why Riker offers no real examples of majority tyranny. Only in an election in which the “issue dimensions are restricted” (Riker, 1982, p. 241) to allow citizens to choose between two binary issues, candidates, and the like, can populist democracy be considered viable. However, as Riker asserts, the manipulation of the political agenda necessary to reduce the voting process to a binary contest has already undermined the populist process. In other words, someone else—an elite of some sort—has stepped in and selected the two issues or candidates from a wider array.

Riker's (1982) rational choice approach drives him to the conclusion that populism is not dangerous, and neither is it tenable. But Riker does not give up on democracy altogether. Can liberalism avoid Arrow's paradox? The answer appears to be yes if by democracy we mean populist democracy, but the answer would appear to be no if by democracy we mean liberal democracy. Riker solves the problems of democracy raised by social choice theory by defining democracy as either liberal or populist, and then demonstrating how social choice theory renders only liberalism defensible. Riker dismisses the populist vision of democracy as “empty,” inconsistent,” and “absurd,” while praising the liberal version as the only kind of democracy that is practically possible. Liberal democracy is viable because, unlike populism,

liberalism requires voters only to choose between two competing issues, candidates, or officials.

Gerry Mackie (2003) has recently written the most thorough and well-argued defense of democracy against the Rikerian social choice critique. Mackie argues that Riker mistakenly assumes that various enduring problems of modern democracy are insurmountable. Social choice theorists such as Riker unfairly stress the *logical possibility* of voting cycles, strategic voting, agenda control, and the other problems associated with aggregating preferences, rather than their *empirical probability*. In other words, Riker has based his defense of liberalism against populism on the theoretical and logical possibility that voting cycles, strategic voting, and agenda control will in fact occur. Riker cites a number of examples drawn from U.S. political history in an attempt to show that the logical possibility of voting cycles, strategic voting, and agenda control have actually occurred. Mackie defends populist democracy by carefully examining the historical examples Riker cites in support of the social choice critique. He argues that Riker's evidence, when carefully examined, does not support his conclusion that populist democracy is impossible. Mackie demonstrates that at key points in Riker's argument, logical and factual errors are made that skew the results in favor of the social choice critique.

Deliberative Theory

The *theory of deliberative democracy* emerged in the 1980s and matured in the 1990s. Deliberative democracy is seen by many political theorists as a replacement for the overly utopian participatory theory and as a response to the rise of the liberal minimalist theory of democracy. Jürgen Habermas (1984, 1987, 1989, 1996) is considered the philosophical father of deliberative democracy and is acknowledged as the thinker most responsible for reviving the emphasis on deliberation. Specifically, Habermas attempts to develop a *theory of communicative action*, the objective of which is to enable citizens to reach consensus over their shared reality and the norms that guide their actions. Ideally, communicative action would produce mutual understanding among citizens about their views of the world. The value of this process ought to be obvious to any student of democratic theory. Seeking to couch his theory in universal terms, he argues for a "transcendental-pragmatic" (Habermas, 1990, p. 130) method. He argues that this approach will free communicative action from any dependency on any specific culture or moral intuitions rooted therein. The transcendental-pragmatic approach is then used to justify a "discourse theory of ethics" whereby the "ideal speech situation" (Habermas, 1990, p. 130) is theorized in which citizens are assumed to be willing and able to deliberate according to the rules inherent in language, and political judgments are made on the grounds that all participants in the deliberative process would, if they could, accept them.

On close inspection, it becomes clear that the central issues of the deliberative theory correspond to the core

issues of the liberal minimalist theory. These core issues have to do with consensus over a common good, the aggregation of preferences, the source and malleability of those preferences, and the source(s) of legitimation of political procedures and of the political system as a whole. At the heart of the liberal minimalist theory is a conception of citizens as passive consumers who exercise control over their leaders by voting, and it understands the political process as a battle for power between competing interests instead of an effort toward a common good. The deliberative model is designed to encourage citizens to actively participate in the political process by seeking consensus about the common good in a public forum. Another assumption of the liberal minimalist theory is that citizens vote according to their preferences. Thus, the primary purpose of a political system is to aggregate preferences through some kind of voting procedure. But the question as to the origins of voters' preferences, how they are formed, and whether they are fixed or mutable is often ignored by liberal minimalists. Deliberative democrats cite this as a fault in the liberal minimalist theory. To counter that defect, they theorize about the process by which preferences are formed. Deliberative theory assumes that a common good can be identified through citizen deliberation. Consequently, when citizens do register their preferences in the voting booth, the legislation that results can be seen as more than a mere registration of individual preferences and rather as a recognition and affirmation of a common good to which citizens each sublimate their private wills.

Habermas's theory of communicative action is central to the theory of deliberative democracy, but it is also a problematic theory. The contentious assumption at its core is the claim that consensus among persons with divergent worldviews can achieve consensus on fundamental social and political norms. Such consensus requires the identification and justification of various *universal* moral principles that participants will freely consent to live by. These principles include the assumption that all citizens are free, equal, willing, and able interlocutors who are actively seeking consensus regarding social and political norms. Habermas derives these principles from his presuppositions about human language and communication. Thus, he argues that human beings implicitly consent to those norms when we seek to understand one another. It is here that Habermas's theory runs into difficulties that continue to beleaguer deliberative democrats. The question remains, on what grounds can we assume that citizens will agree on those critical social and political norms?

Another persistent weakness of Habermas's theory of deliberative democracy is inattentiveness to institutional application. Indeed Habermas recognizes this too, and thus *Between Facts and Norms* (1996) is, at its core, an attempt to show that the theory of communicative action and the theory of discourse ethics are inherent in human reason and language and thus provide deliberative democratic institutions with a transcendental-pragmatic foundation. Yet Habermas fails to articulate a persuasive plan

for the institutionalization of deliberative democracy or to identify a sociological dynamic that makes deliberative democracy a historical possibility. In fairness to Habermas, institutional embodiments of his democratic theories are occasionally explicitly addressed. Even so, Habermas's recent democratic theorizing offers at best only an attenuated vision of a deliberative democratic polity. Moreover, his work has done more to shift the debate in democratic theory further *away* from democratic practice, which is ironic, given his theories' concern with democratic practice.

One of Habermas's goals is to offer an irrefutable philosophical justification of democracy. Habermas argues that the principles of democracy are inherent in the rational structure of language and thus emphasizes the identification of philosophical foundations above the identification of a sociocultural, historical dynamic that would enable deliberative democracy. In so arguing, Habermas strives to situate democratic principles beyond public debate. This is problematic. Habermas theorizes a model of democracy in which debate about social norms is at the very core. Nevertheless, it appears that the norm of rationality and its linguistic expression via democratic deliberation are beyond debate. Indeed, as Habermas asserts, language itself demonstrates, in an *a priori* way, that such principles of democracy are transcendentally given. Hence, Habermas is unwilling to subject some truth claims to debate—the claims of neo-Nazis, for example—because they violate the democratic principle. If Habermas were a democratic pragmatist in the tradition of John Dewey or Richard Rorty, he could not in good faith make such exceptions. But because Habermas does, he is compelled to justify his exceptions, with questionable success.

Agonistic Pluralist Theory

Mouffe (Laclau & Mouffe, 1985; Mouffe, 2000, 2005) is a radical democratic theorist and ardent critic of liberal minimalist and deliberative models of democracy. She is motivated by the belief that older versions of pluralism and of participatory, liberal minimalist, and deliberative theories of democracy are inadequate substitutes for the radical theories of democracy associated with Marxist and social democratic projects. Mouffe refers to her alternative to these theories as *agonistic pluralism*. Mouffe's style of theorizing is in the tradition of Marxist, socialist, and neo-Marxist critical theory, but it has also been shaped by postmodern and poststructuralist thinkers. Most significant among these are Foucault, Derrida, and Wittgenstein. While agonistic pluralism is still neither as developed nor as influential as the liberal minimalist or deliberative theories, it has become increasingly influential during the first decade of the 21st century.

Theorists of agonistic democracy understand politics to be driven by conflict and not consensus. Their emphasis, then, is on encouraging and preserving social diversity rather than attempting to sublimate or eliminate various

identities by means of the illusion of a common good or general will. They are not reformist but instead seek the radical democratic transformation of society through the formation of “counter hegemony.” Because the agonistic model of democracy draws on the postmodern model of society, agonistic democrats view politics as a sphere of intersections between myriad identity, lifestyle, and cultural discourses. All of these represent a challenge to the hegemonic discourse of neoliberalism, by which Mouffe means a political ideology that affirms the sovereignty of the market and privileges its logic in all spheres of society. Such a challenge ought to be encouraged, for it is the only process by which hegemonic discourses can be challenged and undermined. The objective, however, is not the substitution of one hegemonic discourse for another but the creation of a politics of radical indeterminacy and pluralism. In the radical pluralist public sphere envisioned by agonistic democrats, diverse social, cultural, and political forces coalesce around the values of cultural recognition, direct democracy, and performative resistance to the hegemonic discourse. Such resistance is perpetual and reflects the inherent agonistic character of human nature and of politics. It should not be denied or avoided but recognized as a necessary condition for radical democracy.

Mouffe posits an important conceptual distinction between “the political” and “politics” that helps to elucidate her critique of contemporary liberal democracy and the value of the agonistic theory of democracy. Whereas “politics” refers to a political system and sundry processes associated with politics, “the political” is a defining feature of human beings and of their societies. The concept captures and expresses Mouffe's conviction that antagonism is a defining feature of every human society and therefore of every “political” practice and institution. Mouffe repeatedly acknowledges and employs the “political” as an essentially contested term. She argues that most political thinkers in the modernist Enlightenment tradition have attempted to eliminate this antagonism by appealing to universalistic rationalist principles that produce consensus. For Mouffe, such an attempt is misguided, probably impossible, and certainly dangerous for democratic politics because conflict is the product of identity pluralism, which remains a valuable and an inherent facet of the human condition. Thus any attempt to eliminate conflict from the democratic process can only endanger that pluralism.

Mouffe's critique of liberal democracy, encapsulated in what she refers to as the *democratic paradox*, may be summarized by recounting her critique of Rawls's and Habermas's contributions to deliberative democracy. Mouffe concentrates on two issues that illustrate the weaknesses of deliberative theory. The first shortcoming is its attempt to banish value pluralism—a valuable source of agonism—from the public sphere. This is a prerequisite in achieving Rawls's and Habermas's goal of democratic legitimacy via some form of consensus. Mouffe argues that in attempting to do so, both Rawls and Habermas consign pluralism to a nonpublic sphere to insulate politics from its

consequences. The consequences would be the reduction and/or elimination of conflict and agonism, which Mouffe, as noted, argues is the defining quality of “the political.” The second shortcoming is that both authors attempt to reconcile ancient liberty with modern liberty. From Mouffe’s agonistic perspective, the attempt to reconcile these two liberties denies the paradoxical character of modern democracy and an elemental tension between the logic of democracy and the logic of liberalism.

Agonistic pluralists argue that this tension is ineradicable, nor is it a problem. This move on the part of Habermas and Rawls is merely an ill-fated attempt to insulate politics from the inevitable effects of value pluralism. The objective of agonistic democracy is “the negotiation of that paradox” (Mouffe, 2000, p. 93). Consonantly, the goal of the agonistic democratic theorist is to acknowledge this inherent reality and theorize institutions that will not attempt to eradicate this tension but rather will “tame” it in a way that preserves a healthy element of conflict for democratic politics while controlling what is too destructive. Once “the political” has been acknowledged and democratic politics theorized with its exigencies in mind, the next logical step is to begin to consider how the political will manifest in “politics,” that is, what institutions will serve to tame the agonistic.

Critics have highlighted at least three significant problems with Mouffe’s agonistic pluralism. One criticism is that the legitimacy of an agonistic pluralist polity requires a collective consensus on its fundamental principles. This critique suggests that agonistic pluralism shares the same basic flaw that Mouffe identifies in the deliberative theory of democracy. Another criticism involves Mouffe’s recent argument that parliamentary democracy is the ideal political institutional embodiment for an agonistic pluralist polity. Thus, Mouffe’s endorsement of parliamentary democracy seems oddly status quo. Perhaps the most trenchant critique of Mouffe’s agonistic pluralism is made by Aletta Norval (2007). Although receptive to the postmodernism and post-structuralism that shapes Mouffe’s theory of democracy, Norval argues that Mouffe’s theory is excessively abstract and does not engage with the ordinary practices and grammar of political life. The solution is a modification of agonistic pluralism that Norval refers to as *aversive democracy*. Despite these criticisms, agonistic pluralism will likely solidify its place as a major theory of democracy and thus influential in modern democratic thought.

Implications for Modern Democratic Practice

Despite the tenuous connection between democratic theory and practice, there are examples of praxis. Like the proverbial chicken and egg, it is not always possible to discern whether a particular theory of democracy has in fact influenced subsequent democratic practice. It is likely that

democratic theorists draw on democratic practice to invent and refine their theories at least as often as—and probably more often than—their theories serve as the blueprint for novel forms of democratic practice. Placing the inherent difficulties with such speculations aside, several specific examples illuminate how contemporary democratic theory continues to affect modern democratic practice.

One example is the study of *democratic transitions*. Scholarly interest in understanding the problems associated with the political transition from an authoritarian regime to a constitutional democracy has exploded. After the fall of the Soviet Union in 1989, and more recently after the removal of the authoritarian regime in Iraq, political scientists working in the subfield of democratic transitions frequently draw on modern democratic thought in an attempt to determine what form of constitution should be instituted in a particular country. For the most part, it appears that people in charge of constitution building find the liberal minimalist model most suitable. Although in some instances, political scientists have concurred, the field of modern democratic thought offers multiple alternatives to choose from. In the end, scholars often have very little or no direct control over what theory of democracy guides constitution building. Nevertheless, modern democratic thought enables us to explain and critique regimes that are identified as being democratic.

A second example involves one of the most contentious concepts employed by democratic theorists: participation. Democratic participation can take many forms: legislating, protesting, interest group membership, letter writing, and so on. What one democratic theorist may judge an exemplary act of democratic participation may be dismissed by another as inadequate or perhaps even dangerous. The various theories of democracy examined in this essay suggest that participation remains a key concept in modern democratic thought. Those who envision a democracy with high levels of participation across all spheres of society draw on the ideas of participatory democratic theory to challenge regimes that appear insufficiently democratic or clearly antidemocratic. Is voting a sufficient mode of participation in a democracy, as the liberal minimalists argue? Or, as the deliberative democrats would have it, are deliberation, discussion, and debate necessary for democracy? Then again, perhaps the participatory democrats have it right when they argue for an expansive theory of participation that includes the workplace and the household. The concept of participation as elaborated in sundry theories of democracy continues to inspire citizens to justify and criticize their democracy—as well as prompting them to act.

A third example of democratic theory influencing democratic practice is the *deliberative polling technique* developed by James Fishkin (1991). Deliberative polling, as the name implies, derives from the theory of deliberative democracy formulated by Habermas. In the early 1990s, Fishkin proposed that citizens be randomly selected from the population and assembled to hear experts speak

about soon-to-be-voted-on issues. Participants would then discuss and debate among themselves before casting their vote. Evidence supports the claims of deliberative democrats that the deliberative procedure helps citizens formulate preferences that aim at the common good. Moreover, because the sample of citizens is representative, we can generalize from the particular to the general and conclude that how the sample size votes would be how all citizens would vote—if they had a similar deliberative opportunity. Recently Fishkin and Bruce Ackerman (2005) have proposed the idea of “deliberation days.” The authors recommend that citizens be provided at least 1 day a year during which they fulfill the deliberative democratic ideal: essentially a holiday for democratic deliberation. Although these proposals are relatively experimental, they hold great promise for modern democracy.

Directions for Future Research

There is no shortage of research projects in democratic theory awaiting future graduate students. Here is a selection of broad research topics that political scientists should explore further during the next decade.

One topic for future research involves the pluralist theory of democracy. The question that political scientists should be investigating is to what extent theory accurately describes modern democracies. Whereas the criticisms of the neopluralists were, for the most part, integrated into the theory of pluralism, the advocates of the “new pluralism” have argued that the theory of pluralism must be further modified in light of new philosophical insights. Just how that might be done is a line of inquiry that needs to be pursued. Second, the growing number of participatory institutions in Brazil and elsewhere around the globe portends a revival of participatory democratic theory. That theory can then be modified in light of the wealth of new data now available on participatory democracy. Third, some cursory research has been done regarding the potential synthesis of deliberative democracy and liberal minimalist democracy. More work needs to be done to highlight how deliberative procedures might engender the sort of public interest voting that minimalist theorists of democracy argue is not possible. Fourth, although a fair amount of research on deliberative polling has been conducted, a detailed investigation into its potential to deepen democratic practice is still wanting. Moreover, just what those mechanisms and institutions would look like needs further articulation (e.g., deliberative polling, deliberation days, and citizen juries).

Indeed these mechanisms designed to enhance (deliberative) democracy are but a few of many imaginative ideas that suggest a fifth area of research involving various democratic innovations, including participation budgeting, minipublics, direct legislation, e-democracy, and so on. And while such innovations often garner the most attention, the older and often less provocative idea of political

representation—a sixth area for future research—is currently experiencing a renaissance that will likely gain in momentum over the next several decades. Finally, although agonistic democracy has been sufficiently theorized, the sorts of institutions that might be associated with that theory have yet to be articulated. Whereas a parliamentary form has been suggested by Mouffe, the case has yet to be made as to how agonistic parliamentary institutions transform antagonism into agonism. These are but a few of the numerous possible research topics. In short, there is much work yet to be done.

Conclusion

Democracy continues to be a subject of supreme interest among political scientists. Consequently, the state of contemporary democratic theory is vibrant and will likely remain so indefinitely. It should be obvious now that students interested in pursuing a career as a political scientist (especially in the United States, but internationally as well) ought to keep in mind the discipline’s enduring interest in democracy. Prudent students will make a serious effort to become conversant with modern democratic thought. This entails becoming familiar with the various theories of democracy that have been formulated by political scientists and theorists over the past 50 years and understanding the contexts in which they have developed and the ways in which they conflict and overlap. The five most prominent theories, pluralist, participatory, liberal minimalist, deliberative, and agonistic, all contribute to our understanding of modern democratic thought. Because democracy is at the heart of American political science, one or more of these theories will likely enter into a political scientist’s research. Whether a student chooses to focus on comparative politics, international relations, American politics, public administration, political theory, or one of the relatively new subfields on the ever-expanding list, the topic of democracy will inevitably arise. In the 21st century, the defining characteristic of a competent political scientist undoubtedly will be a comprehensive understanding of modern democratic thought.

References and Further Readings

- Arrow, K. (1962). *Social choice and individual values*. New Haven, CT: Yale University Press.
- Avritzer, L. (2009). *Participatory institutions in democratic Brazil*. Baltimore: Johns Hopkins University Press.
- Barber, B. (1984). *Strong democracy: Participatory politics for a new age*. Berkeley: University of California Press.
- Campbell, D., & Schoolman, M. (2008). *The new pluralism: William Connolly and the contemporary global condition*. Durham, NC: Duke University Press.
- Dahl, R. A. (1956). *A preface to democratic theory*. Chicago: University of Chicago Press.

- Dahl, R. A. (1961). *Who governs? Democracy and power in an American city*. New Haven, CT: Yale University Press.
- Dahl, R. A. (1986). *A preface to economic democracy*. Berkeley: University of California Press.
- Dahl, R. A. (1989). *Democracy and its critics*. New Haven, CT: Yale University Press.
- De Sousa Santos, B. (Ed). (2005). *Democratizing democracy: Beyond the liberal democratic canon*. London: Verso.
- Downs, A. (1956). *An economic theory of democracy*. New York: Harper.
- Fishkin, J. S. (1991). *Democracy and deliberation*. New Haven, CT: Yale University Press.
- Fishkin, J. S., & Ackerman, B. (2005). *Deliberation day*. New Haven, CT: Yale University Press.
- Gould, C. C. (1988). *Rethinking democracy: Freedom and social cooperation in politics, economy, and society*. Cambridge, UK: Cambridge University Press.
- Gunnell, J. G. (2004). *Imagining the American polity: Political science and the discourse of democracy*. University Park, PA: Pennsylvania University Press.
- Habermas, J. (1984, 1987). *The theory of communicative action* (Vols. 1-2). Boston: Beacon Press.
- Habermas, J. (1989). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge: MIT Press.
- Habermas, J. (1990). *Moral consciousness and communicative action*. Cambridge: MIT Press.
- Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy*. Cambridge: MIT Press.
- Hamilton, A., Jay, J., & Madison, J. (1966). *The federalist papers*. New York: Doubleday.
- Held, D. (2006). *Models of democracy* (3rd ed.). Stanford, CA: Stanford University Press.
- Laclau, E., & Mouffe, C. (1985). *Hegemony and socialist strategy: Towards a radical democratic politics*. London: Verso.
- Leonardo, A. (2009). *Participatory institutions in Brazil*. Baltimore: Johns Hopkins University Press.
- Lindblom, C. E. (1977). *Politics and markets*. New York: Basic Books.
- Mackie, G. (2003). *Democracy defended*. Cambridge, UK: Cambridge University Press.
- Macpherson, C. B. (1977). *The life and times of liberal democracy*. Oxford, UK: Oxford University Press.
- McGann, A. (2006). *The logic of democracy: Reconciling equality, deliberation, and minority protection*. Ann Arbor: University of Michigan Press.
- Mouffe, C. (2000). *The democratic paradox*. London: Verso.
- Mouffe, C. (2005). *On the political*. London: Routledge.
- Norval, A. (2007). *Aversive democracy: Inheritance and originality in the democratic tradition*. Cambridge, UK: Cambridge University Press.
- Olsen, M. (1971). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Pateman, C. (1971). *Participation and democratic theory*. Cambridge, UK: Cambridge University Press.
- Rawls, J. (1993). *Political liberalism*. New York: Columbia University Press.
- Riker, W.H. (1982). *Liberalism against populism: A confrontation between the theory of democracy and the theory of social choice*. San Francisco: Freeman.
- Shapiro, I. (2003). *The state of democratic theory*. Princeton, NJ: Princeton University Press.
- Shapiro, I., Stokes, S. C., Wood, E. J., & Kirshner, A. S. (2009). *Political representation*. Cambridge, UK: Cambridge University Press.
- Smith, G. (2009). *Democratic innovations: Designing institutions for citizen participation*. Cambridge, UK: Cambridge University Press.
- Urbiniati, N., & Warren, M. (2008). The concept of representation in contemporary democratic theory. *Annual Review of Political Science*, 11, 387-412.

72

MODERN LIBERALISM, CONSERVATISM, AND LIBERTARIANISM

JOHNNY GOLDFINGER

Marian University

Liberalism and conservatism were the original rivals in modern Western political theory beginning in the latter half of the 18th century. Since then, there have been many important theoretical developments. For example, other important political theories have emerged, most notably socialism, and the liberal tradition has branched into two competing wings, modern liberalism and libertarianism. Even with these changes, however, the debate between liberalism and conservatism remains a fundamental feature of political theory and practice. This chapter provides an introduction to modern liberalism, libertarianism, and conservatism. It discusses the basic features of these political theories, focusing on their views of human nature and understandings of fundamental political values. It also provides a survey of important classic and contemporary works of modern liberal, libertarian, and conservative political thought. Last, this chapter discusses the perennial questions that characterize these theories and examines their future prospects.

Human Nature and Political Values

Political theory is the study of how society ought to be organized and regulated. Thus, it is essentially about the way people ought to live together. Political theory examines the nature of political institutions, the responsibilities

of the state, and the social norms and expectations that shape the way people think and behave. It is concerned with both the normative implications and practical consequences of social and political institutions. Therefore, it takes into account what is desirable and necessary. Political theory creates a model of politics that serves as “a basis for judging the wisdom of our political deeds and arrangements” and provides a “perspective on what we are doing and thereby some grounds for seeing how well we are doing it” (Spragens, 1975, p. 5).

Leading contemporary political theories include modern liberalism, libertarianism, and conservatism. A common way to distinguish between these three competing theories is by their understandings of the state’s practical responsibilities. Modern liberals, for example, advocate an activist state that looks after the welfare of its citizens. It should provide the basic resources people need to pursue their interests effectively. Libertarians argue for a state with very limited powers and responsibilities. It favors minimal government intervention in economic, social, and personal affairs. Conservatives believe the state should help maintain social order and reflect each society’s unique traditions and customs. They typically favor the status quo and are suspicious of radical change.

The basic political institutions and policies that these theories espouse are functions of their views of human nature and the political values they prioritize. All political

theories are based on views of human nature. "To do otherwise would be to take the complex and perhaps unpredictable human element out of politics" (Heywood, 1999, p. 16). Human nature refers to human qualities and potentials such as rationality, cooperativeness, benevolence, dependability, fidelity, and so forth. Political theories make assumptions about what people are capable of doing. A political theory's view of human nature is essential for the success of its institutions and policies. If a theory is too optimistic about human nature, then the state will not provide enough control and structure for its citizens. If a theory's expectations are too negative, then the state will be overly paternalistic and not allow people to develop fully as human beings.

Political values are the basic concepts and principles that people believe are important. They include a range of ideals such as autonomy, liberty, equality, justice, virtue, and social stability. Political theories make claims about what values are politically significant and the way they should be prioritized. These political values are fundamental for political theory because they represent the ultimate purpose of the state. The aim of political theory is to advance these values by constructing appropriate political institutions and policies.

Human nature can be understood as the practical limits of politics. It determines what is possible. Political values, in contrast, are the theoretical goals. They represent the aspirations for how people want to live and arrange their collective existence. Political theory involves creating and justifying political institutions and policies that advance political values in a way that is consistent with human nature. Therefore, to better understand modern liberalism, conservatism, and libertarianism, it is useful to examine their views of human nature, fundamental political values, and rationalizations for particular political institutions and policies.

Modern Liberalism and Libertarianism

Although modern liberalism and libertarianism represent distinct political theories, they are both part of the liberal tradition and, thus, share similar liberal assumptions and commitments. The liberal tradition's primary political concern is the interest of the individual. Its fundamental political values include liberty, equality, autonomy, consent, and toleration (see Kelly, 2005; Raz, 1986). Liberty is particularly important for the liberal tradition, as the term *liberal* suggests. It considers the promotion of individual freedom a fundamental responsibility of the state. This emphasis on freedom makes toleration a necessary value for the state and society (see Gray, 2000). The liberal tradition is committed to the value of equality as well, recognizing the equal moral worth of all citizens (see Gutmann, 1980). It also respects the autonomy of each individual. The state, therefore, should allow people to act as independent moral

agents. It follows that consent is a basic component of legitimacy. Because individuals are independent moral agents, their consent creates a binding moral obligation (see Plamenatz, 1938).

The liberal tradition's emphasis on individual freedom and autonomy means that people must take responsibility for their own choices. The ability to make personal decisions, however, is useful only if individuals are able to handle the responsibility. The liberal tradition, therefore, is marked by an optimistic view of human nature. It assumes, for example, that individuals are capable of rational thought, as well as judging what is in their own best interests, cooperating with others, and being trustworthy. These qualities allow humans to act on their freedoms and autonomy without undermining the well-being of society.

Modern liberalism and libertarianism share the values of the liberal tradition and its view of human nature. They differ significantly, however, in their understandings of two key concepts, liberty and equality. There are two distinct ways to interpret the meaning of liberty (see Berlin, 2002). *Negative liberty* refers to the absence of restraint or impediments. People are free, in this negative sense, as long as they can act without interference by other human beings. *Positive liberty* involves being able to act and fulfill one's desires. Simply having negative liberty does not mean people will be able to take advantage of it in practice. Positive liberty is having the resources and power to take action, live a certain way, or achieve some sort of self-realization (see Green, 1986). Thus, intervention by the state may be required in order to give people what they need to satisfy their aspirations.

Equality also has different interpretations (see Pojman & Westmoreland, 1996). It has a formal meaning, which simply involves people being considered morally equal and treated equally under the law. *Formal equality* essentially removes barriers to action. Equality can also be understood in terms of opportunity (see Roemer, 1998). *Equal opportunity* focuses on giving people a fair chance to compete and succeed. It attempts to remove any normatively unacceptable disadvantages. This type of equality is concerned only with the quality of competition or the process. *Equality of outcome*, in contrast, is interested in the end result (see Rawls, 1971). It is committed to the notion of *redistribution*, so that people end up with the same goods, resources, or services.

Modern liberalism adopts a positive understanding of liberty and prioritizes equal opportunity and, in some cases, equality of outcome. The state is expected to provide resources and opportunities that allow people to act more effectively and make competition more fair and equitable. It should distribute goods and provide services for the purpose of expanding individual choices and improving opportunities to compete successfully. Thus, modern liberals advocate an activist state. Libertarianism differs in that it is committed to negative liberty and formal equality. The primary function of the state should be to prevent

illegitimate uses of force, thereby protecting people, their property, and their ability to act without interference. Otherwise, individuals should be allowed to make their own decisions and look out for their own well-being without the state's intervention. Formal equality simply requires that the government treat all citizens equally. Libertarians, therefore, favor a state with limited powers and duties.

Origins of the Liberal Tradition

Modern liberalism and libertarianism differ greatly in their understandings of the responsibilities of the state, yet they both trace their roots to the liberal ideals that emerged during the 17th century in Europe. Liberalism was the key political innovation of the early modern period. During the Middle Ages, social and political life was characterized by inequality and submission. The feudal system, for example, was an important economic and social institution that maintained a rigid hierarchy through ascribed social status. The Catholic Church was the dominant religious institution. Its enforcement of religious conformity secured the authority of Scholastic understandings of morality and political order. A series of events, however, led to the end of feudalism and weakened the Catholic Church's influence. Development of new productive technologies and discoveries of new lands and trade routes contributed to the demise of feudalism. So did the plagues, which undermined the social hierarchy by leaving gaps in the noble class. The Protestant Reformation challenged the doctrines of the Catholic Church, placing special emphasis on the spiritual equality of all individuals. The Renaissance reestablished a humanistic perspective, focusing attention back on human concerns as opposed to God and spiritual matters. The modern scientific revolution, which called into question Scholastic science, undermined the authority of Scholasticism in general.

These changes inspired a new theory of politics, which came to be known as liberalism. Liberalism's emphasis on liberty and equality reflected the challenges to the medieval norms of ascribed status and religious conformity. The focus of liberalism was the individual. It viewed the function of the state as serving the interests of each citizen, not God, a king, or the aristocracy. Liberalism also made *consent* the basis for political legitimacy. The authority of the state was justified through the consent of the people, not divine right, tradition, or force.

The main forerunner of liberalism was the English philosopher Thomas Hobbes (1588–1679). Although he falls short of articulating a liberal political theory, the ideas he presents in *Leviathan* (1651/1994) reflect many of the basic assumptions and priorities of the liberal tradition. Hobbes's political theory attempts to understand the origins of the state in terms of individual self-interest. He begins by imagining a condition in which there is no state and people have complete freedom to act as they wish.

Hobbes contends this condition would be so miserable that individuals would choose to give up their complete freedom and voluntarily submit themselves to the absolute authority of a political sovereign for protection. The state gains its legitimacy through the consent of the individual.

Hobbes (1651/1994) assumes that, in the state of nature, all men are naturally free and equal with an impulse to pursue one desire after another. He also contends that human beings are naturally competitive, diffident, and vain. Conflict, therefore, is a by-product of human nature. It follows that the state of nature is actually a "state of war" (p. 76) with "every man against every man" (p. 76). He famously concludes that life in the state of nature is "solitary, poor, nasty, brutish, and short" (p. 76). People, however, are rational enough to figure out a way to leave the state of nature. Hobbes claims that individuals will create and consent to a social contract that establishes an all-powerful sovereign who is strong enough to protect them from each other and external threats. This authoritarian government is contrary to the liberal commitment to personal freedom and autonomy. So, in this fundamental sense, Hobbes is not a liberal. However, his assumptions about the importance of the individual, the consensual basis of political legitimacy, and the naturalness of human freedom and equality anticipate key liberal ideals. It is Hobbes's low view of human nature that compels him to sacrifice the individual's natural liberty for personal safety and, thus, prevents him from being a liberal political theorist.

The person most responsible for bringing liberalism to the intellectual forefront is the English philosopher John Locke (1632–1704). His *Two Treatises of Government* (1689/1988) was and continues to be an extraordinarily influential work. Locke adopts several of Hobbes's basic ideas, including the concept of a state of nature and the social contract. His assumptions about human nature, however, lead him in a different political direction. Locke starts with people in a state of nature. He claims there are laws of nature, including the prohibition of individuals from harming others in their "life, health, liberty, or possessions" (p. 271). Humans are generally rational enough to comprehend these laws of nature through their reason and moral enough to obey them. If a person violates the laws of nature, then others have the right to enforce the laws and punish offenders in proportion to the crime. Locke, however, assumes that humans cannot remain objective when their own interests are at stake. Consequently, when someone violates the law of nature, the victims or their family and friends are likely to overpunish the offender. In doing so, however, these people are themselves violating the laws of nature and, thus, subject to punishment. The result is a vicious circle of overpunishment and the violation of the laws of nature. Locke claims that this condition makes the state of nature "inconvenient." To avoid this situation, rational people will contract to set up a state to act as a mediator to settle disputes between people. The state effectively serves as an impartial judge, jury, and executioner. Thus the primary function of

the state is to objectively enforce the law of nature, allowing citizens to enjoy their life, liberty, health, and possessions without interference by others. If the state is unable or unwilling to fulfill its obligations, then the contract is broken and citizens have the right to withdraw their consent and revolt if necessary.

Negative liberty is a primary concern for Locke's social contract theory. Thus, he is often cited as an inspiration by many libertarian thinkers. Locke has also had an influence on liberal thought more generally through his advocacy of tolerance in *A Letter Concerning Toleration* (1689/1983). In this essay, he distinguishes between a public and a private realm, arguing that religion is properly a private matter that does not have a bearing on the well-being of the state. Hence, there is no reason to impose religious conformity. As long as a religion does not affect the order of the state, it should be tolerated. This defense of religious toleration reinforces liberalism's commitment to freedom and personal autonomy. Locke's distinction between public and private matters has also been applied more broadly by liberals to argue for expanded civil liberties and less government regulation of lifestyle choices.

Libertarian Theories of Politics

Libertarianism advocates negative liberty and formal equality in all aspects of society, including the economic and private spheres. Libertarians support laissez-faire economic policies, including a free market and the protection of personal property rights. They also favor a laissez-faire approach to personal behavior. All people should be able to do and think as they please provided that they do not harm others.

Libertarian economic ideals are reflected in the work of Adam Smith (1723–1790), the Scottish economist who wrote the seminal defense of capitalism in *Wealth of Nations* (1776/1976). Smith applies liberal concepts such as freedom, equality, consent, individuals, and rationality to argue for the establishment of a free market. A free market produces economic outcomes that benefit society and gives people opportunities to advance themselves through their own merit. Allowing people to act freely in the marketplace tends to keep prices relatively low and leads to the development of new and better products. What makes this outcome possible is the natural self-interest of human beings. Allowing people to pursue their own self-interests in the marketplace creates the competition that produces economic benefits for the consumer and conditions for producers to advance on merit. If the state interferes in the market and undermines this competition, then it also undermines the benefits. Thus, Smith was generally opposed to the regulation of the market by the state.

The libertarian concern for personal freedom in private affairs is perhaps best articulated by the British philosopher John Stuart Mill (1806–1873) in *On Liberty* (1859/1978). Mill, who is best described as a civil libertarian,

argues that human beings should be allowed wide latitude in behavior and thought, with limits justified only when others are injured. He called this criterion the *harm principle*. Mill's justification for freedom from government interference in the individual's private life is based on the utility it produces for society. He claims that freedom of speech and thought facilitates the pursuit of truth and elimination of error. This freedom gives people a chance to challenge accepted ideas and norms. It, therefore, provides opportunities for the truth or falsity of the status quo to be determined. Freedom of action also benefits society. Mill claims that the ability to pursue different lifestyles creates "experiments in living" that people can observe and from which they can learn. Thus, it allows people to make better informed decisions about which lifestyle they want to adopt. Mill's association of freedom with progress reflects his faith in the ability of human beings to make decisions for themselves.

Ayn Rand (1905–1982), a Russian who immigrated to the United States in 1926, is most responsible for bringing libertarianism to the attention of the general public. She was a successful novelist turned philosopher who developed the philosophy of *objectivism*, which espoused many libertarian ideals. Rand is perhaps best known for her novel *Atlas Shrugged* (1957/2005), which warns of the dangers of excessive economic regulation by the government and celebrates productivity, rational selfishness, and a free market economy. These themes are also reflected in the ethical and political dimensions of Rand's objectivist philosophy. She contends that the greatest human virtue is the exercise of reason and the highest moral purpose in life is the rational pursuit of self-interest. People should live only for themselves and not sacrifice for others. Thus, the only function of government should be to protect individuals from coercion so they can use their reason and pursue their own interests without interference. For similar reasons, Rand also advocated laissez-faire capitalism.

The *Austrian School of economics* has had a particularly significant influence on contemporary libertarian thought through the works of economists such as Ludwig von Mises (1949/2007) and Friedrich Hayek (1960/1978). The Austrian School advocates a laissez-faire economy, with the state's main purpose being the strict enforcement of property rights and contracts. Murray Rothbard (1926–1995), an American economist, applies the ideals of the Austrian School to directly advance libertarianism. Rothbard uses the concept of self-ownership and natural rights to defend the value of liberty in *Ethics of Liberty* (1982/2003). He contends it is self-evident that people have ownership in themselves and the goods they appropriate or create. This ownership implies that people should be able to dispose of their property as they wish without interference from the state. Rothbard goes on to examine and critique various types of government coercion in *Power and Market* (1907/2007), concluding that the state is neither desirable nor useful.

Perhaps the most notable contemporary argument for libertarianism was written by the American philosopher Robert Nozick (1938–2002). In *Anarchy, State, and Utopia* (1974), he defends the idea of a “night watchman” state. He begins with the idea of self-ownership; people own themselves, including their body, talents, labor, and whatever they produce. Hence, they have rights to self-determination in themselves and their property that act as “side-constraints,” protecting them from interference by other individuals and institutions such as the state. Nozick does, however, believe that coercive institutions and relationships are legitimate if people give their consent. He believes, for example, that a minimal state—which protects against force, theft, and fraud, enforces contracts, and administers justice—is a practical necessity. People, therefore, would willingly pay to have this protection. Any additional powers of the state to regulate the economy, society, or morality are legitimate only if the people give their consent.

Libertarians use a number of different rationales and justifications to defend their theories, including natural rights, efficiency, utility, progress, the rational nature of humans, and self-ownership. Thus, there are many different versions of libertarianism. Regardless, they all aim for a state that has minimal responsibilities, promotes negative liberty, and treats citizens equally. Libertarianism often focuses on economic freedom. However, it is important to keep in mind that this philosophy favors negative liberty in all aspects of life as long as one’s actions do not harm the freedoms of other people.

Modern Liberal Theories of Politics

Modern liberalism focuses on individual welfare. It developed as part of the reform movement to improve the conditions of the less fortunate through state services and the redistribution of resources. Modern liberals differ in the policies they propose. However, they all have a common interest in using the state to help people pursue their interests and take advantage of their freedoms.

Modern liberalism’s ideals are reflected in the thought of Thomas Paine (1737–1809), one of the Founding Fathers of the United States. He was not only a revolutionary but a social reformer as well. The influence of the liberal tradition is reflected in the slogan attributed to him, “Give me liberty, or give me death.” It is reflected as well in Paine’s pamphlet *Common Sense* (1776/1982). In it, he assumes a state of nature in which people are naturally free and equal, arguing that the purpose of the state is to provide security for the individual and that the best form of government is a representative democracy. Paine was also an advocate for social reform and policies that advanced positive liberty. In his *Rights of Man, Part the Second, Combining Principle and Practice* (1791, 1792/1992), he details extensive social reforms that states should enact. Paine recommends a number of governmental policies

intended to improve the condition of the poor. They include free public education; guaranteed employment; public assistance for the poor, with supplemental support for children, widows, and the elderly; tax relief for the lower class; and a heavily progressive tax on the aristocracy. Thus, Paine believed the state should play an active role in providing positive liberty for its citizens.

Thomas Hill Green (1836–1882), an English philosopher, advocated similar policies to advance positive liberty. He contends that simply having negative liberty is of no value if people cannot take advantage of it. Freedom is not valuable simply for the sake of freedom but because it allows people to pursue their interests and develop as human beings. If people do not have the resources that allow them to put their freedom to use, then the freedom they have is simply an abstraction. The state, Green (1986) argues, should ensure that people have access to resources such as education, health care, housing, and food. Educated, healthy people, who do not have to worry about their subsistence needs, are better able to act on the freedoms they have. Green, moreover, claims that the guarantee of these resources benefits everyone, including the well-off, because they no longer have to worry about these concerns. Hence, the intervention of the state promotes the welfare of citizens by allowing them the opportunity to better pursue their interests and self-development.

Another important voice for modern liberalism was Herbert Croly (1869–1930), a leader of the American progressive movement at the turn of the 20th century. He was particularly troubled by laissez-faire individualism and the industrial revolution, which produced extreme inequalities in wealth and power. These morally and socially undesirable inequalities produced economic and social abuses that were corrupting the country. The competitive pursuit of money was also undermining genuine individuality. Croly (1909/1989) believed a stronger, centralized federal government was needed to effectively address these problems. He favored a state that distributed a share of the wealth and benefits to the whole community and looked out for the welfare of its citizens. The state, at the very least, should ensure that people have some degree of economic power and responsibility and make poverty a negligible social factor. Moreover, it should intervene to raise the laborers’ standard of living and increase their economic independence.

John Dewey (1859–1952) was an American philosopher and educational reformer critical of the individualism associated with negative liberty. He believed that a commitment to liberty required the recognition of the intimate relationship between the individual and society. The individual should not be considered apart from society. Dewey (1930/1999) claims that each person is a product of social, economic, and political institutions. An abstract understanding of the individual misses the real nature of humans. It is also misleading because it seems to suggest

that negative liberty is more consistent with what it means to be an individual, that is, distinct from society. Dewey argues, however, that the social nature of humans means liberty should be understood in a positive sense. Freedom involves being able to shape the social conditions that shape the individual. It allows humans to be “individualized” selves. Thus, freedom is valuable as an activity in which there is a collective exercise of social and self-development.

Perhaps the most influential modern liberal is the American philosopher John Rawls (1921–2002). His *Theory of Justice* (Rawls, 1971) provides a seminal argument for the promotion of positive liberty by the state. He develops a modified version of the social contract that requires people to choose principles of justice in an “original position.” People in this original position are behind a veil of ignorance, which prevents them from knowing anything about themselves. The veil of ignorance is intended to ensure that people choose principles of justice objectively. They are unable to choose principles that advance their own interests because they do not know what their interests are. Rawls (1971) claims that people in the original position would choose two lexically ordered principles of justice. The first principle of justice is “each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all” (p. 302). The second principle of justice is “social and economic inequalities are to be arranged so that they are both (a) to the greatest advantage of the least advantaged, and (b) attached to offices and positions under conditions of fair equality and opportunity” (p. 302). Thus, Rawls contends that the state should protect personal liberties while redistributing primary goods so that the disadvantaged can have the resources and opportunities to live good lives.

Modern liberals believe the state needs to intervene to provide citizens with certain resources so they can act on the freedom they have. The nature and purpose of this intervention, however, can differ greatly. Modern liberals propose a wide variety of government-provided services, programs, goods, and opportunities for the general public and the needy in particular. Modern liberals also have a range of aims such as the elimination of poverty, ensuring the welfare of the least advantaged, promoting equal status, and facilitating self-development.

Modern liberalism’s commitment to positive liberty and equal opportunity leads to the institution of a *welfare state*, that is, a state actively involved in securing the welfare of its citizens. Modern liberals may also adopt socialist means to realize its goals. Welfarism, however, should not be confused with *socialism*, which advocates state control over the means of production and distribution of goods. Welfarism and socialism are certainly compatible political theories, and modern liberals often advocate both welfare and socialist policies. Still, they have different implications for the responsibilities of the state.

Conservatism

Conservatism, as the term implies, emphasizes the value of conservation. Its primary concern is to preserve the customs and traditions of society. This orientation is based on the conservative view of human nature, which assumes that humans are inevitably flawed both intellectually and morally. Conservatives, therefore, are suspicious of any social, economic, or political changes in the name of innovation or perfectibility. They are wary of innovation because they do not trust what is new and unproven in practice. Conservatives disapprove of efforts to achieve perfectibility because they do not think it is possible. Instead of leading toward progress, innovation and perfectibility can actually be dangerous because humans have limited ability to understand or appreciate the consequences of abstract ideas.

Custom and tradition have the advantage of being proven over time. For conservatives, “standing the test of time” is the most meaningful criterion for determining worth because individual judgment is suspect. The value of customs and traditions is proven in practice across the generations. If customs and traditions continue to exist successfully, then there must be something valuable or useful about them. Still, it is important to recognize that conservatives are not against change per se. Sometimes it is necessary. However, these changes should typically be slow and deliberate. They should involve reform, not innovation. Because of the limits on human foresight, change should be gradual to minimize the possible danger to society. Through a practice of slow and deliberate change, problems can be recognized before they do too much harm. Conservatism is a cautious ideology that is sympathetic to the maxim that “a bird in the hand is worth two in the bush.”

Conservatism is also skeptical of political abstraction and universalism. Each society has its own set of historical circumstances that makes it unique. Thus, the appropriate political institutions and policies for a society are a function of its traditions and customs. What is appropriate for one society is not necessarily appropriate for another. Human beings are shaped by their traditions and customs and have limited ability to adapt to new situations.

The precursors of conservatism can be found in pre-modern ideals and principles that emphasized human limits and frailties. The Judeo-Christian tradition, for example, has the concept of pride, which involves excessive self-love and conceit. Adam and Eve’s exalted perception of themselves and what they were due led them to defy God’s command and eat from the tree of knowledge of good and evil. Their sins in the Garden of Eden reflect the human susceptibility to temptation, especially appeals to people’s vanity and the manipulation of their pride. The Christian concept of original sin refers to the inherent moral flaws of human beings. People are sinners by nature. It is part of the human condition. People are all

vulnerable, for instance, to the seven deadly sins: lust, gluttony, greed, sloth, wrath, envy, and pride.

The conservative intellectual tradition took shape in Europe during the 18th century as a response to the spread of liberalism. It argued for the maintenance or return to more aristocratic social and political institutions, which were challenged and undermined by liberal reformers. Thus, conservatism began as a defense of the aristocratic political order, but it has come to involve much more than that. Conservatism is a rich and varied tradition.

Edmund Burke

The Irish philosopher and statesman Edmund Burke (1729–1797) is widely regarded as the father of modern conservatism. His understanding of conservatism is articulated most clearly in response to the dangers he saw in the French Revolution. Writing at the outset of the Revolution in 1789, Burke (1790/1987) warned that the challenge to the French monarchy was a reckless and hazardous undertaking. He was particularly critical of the Enlightenment ideals, such as the faith in human reason and reliance on abstract rationality, that influenced the French Revolution. Burke believed the revolution's abstract principles such as liberty, equality, and the "Rights of Man" were a precarious foundation because they ignored the practical political realities of France. It had a tradition and history of monarchical government and no experience with democratic institutions. The French Revolution tried to impose popular government on a people who were not prepared for the responsibility. Burke's skepticism about the ability of the French to adjust successfully to a new political system reflects the conservative view of human nature. When people are placed in unfamiliar situations without preparation, it is unrealistic to expect them to adapt quickly or behave appropriately. Giving people political freedom without any experience in managing it is to invite abuse and excess. Freedom can be valuable. However, it must be properly constrained and managed.

Burke also found the French Revolution dangerous because it brought sweeping social and political changes to France. In so doing, it ripped apart the "social fabric," which is the foundation for social interactions. The social fabric is sewn from a society's traditions, customs, and mores. Its threads consist of practices, beliefs, values, habits, and rituals that have been handed down from generation to generation. Over time, these practices and beliefs are woven together, complementing and reinforcing each other. If they are widely shared and respected, then the result is a strong social fabric that holds a people together and orients their activity. A common set of practices and beliefs allows people to anticipate the actions and reactions of others, thereby allowing them to coordinate more effectively and avoid conflict. If the social fabric is torn or frayed, it reflects the unraveling of important relationships and commonalities that ground social stability. If

the social fabric becomes completely undone, then people are left without any way to orient themselves in society.

Thus, Burke believed customs and traditions should be respected and should be changed only when absolutely necessary. Because of the complexities of society and the limits of human reason, it is difficult for people to predict the consequences of social changes. Hence, Burke favors having a prejudice for the status quo. Unless there is compelling evidence to the contrary, people should give the benefit of the doubt to long-standing traditions and customs.

Even though Burke was opposed to the French Revolution, he was not opposed to revolutions in general. For example, he approved of the Glorious Revolution in England and was sympathetic to the American Revolution. In both cases, the revolutions were fought to regain lost rights and privileges. They were revolutions in the sense of a revolving back to a previous condition. The French Revolution, in contrast, involved the establishment of a new political and social order and was therefore an attempt at innovation, as opposed to reform.

Conservative Theories of Politics

Another important conservative thinker, who was Burke's contemporary, is the Sardinian statesman and philosopher Joseph de Maistre (1753–1821). His political outlook was greatly informed by Christian pessimism and born out of a reaction to the Enlightenment and the French Revolution. Maistre (1809/1959) was greatly influenced by the concept of original sin and the fallen nature of man. He believes human pride, in particular, was the fundamental feature and flaw of the Enlightenment, which was, he claims, a rebellion against traditional authority and had misguided and dangerous aspirations of perfectibility. He has a similar critique of the French Revolution, which he thought was a vain attempt to institute political perfection through human reason and social construction. Maistre believes in the absolute sovereignty of God and his ordained order. Hence, he expresses contempt for concepts like the social contract and individual consent. Political legitimacy is not a function of human choice but of God's will.

The American sociologist Robert Nisbet (1913–1996) is also critical of the consequences of liberalism. In *Quest for Community* (1953/1990), he claims liberalism has produced a sense of anxiety and alienation that is driving people to search for community. This condition is created by the liberal emphasis on freedom, equality, individualism, and free markets. The liberal state and liberal values have weakened and undermined traditional social organizations such as the family, church, and neighborhood associations. These organizations constrain their members with demands and requirements. Liberalism compels people to free themselves from these restrictions. They even turn to the state for help in challenging and diminishing the power

of these institutions. Although these institutions are restrictive, they provide members with moral bearing, personal security, and a sense of meaning and purpose. With the disruption of these local bases of community, people look to the state for the moral, psychological, and economic support they need. Moreover, the destruction of local institutions leaves people susceptible to totalitarian control. These institutions, which are conduits for collective power, serve as buffers between the individual and the state. Without them, people are left as individuals to stand alone against the state's power.

The conservatism of the English philosopher Michael Oakeshott (1901–1990) is based on skepticism about rationalist thought and the abstract political principles and ideals that it produces. The rationalist attempt to create social and political institutions out of abstract ideals undermines the traditions that have proven themselves through practice and experience. Traditions, which are time tested, are preferable to ideals that may promise more and better but have not been tested in practice or context. Oakeshott (1991) eloquently states this perspective as follows:

To be conservative, then, is to prefer the familiar to the unknown, to prefer the tried to the untried, fact to mystery, the actual to the possible, the limited to the unbounded, the near to the distant, the sufficient to the superabundant, the convenient to the perfect, present laughter to utopian bliss. (p. 408)

Oakeshott's skepticism is also reflected in his preferred political institutions. States, according to Oakeshott, can be understood as either enterprise or civil associations. *Enterprise associations* are characterized by an agreement among members to pursue some particular purpose. *Civil associations* simply involve a body of laws that are used to regulate society so people can pursue their own unique goals. Oakeshott opposed any grand political enterprises, favoring instead the civil association model for state and a rule of law that was consistent with the society's traditions and customs.

The conservatism of the political philosopher Leo Strauss (1899–1973), a German-born immigrant to the United States, is rooted in the premodern political tradition that focused on virtue and the pursuit of human excellence. Strauss (1953/1999) is particularly critical of liberalism because of what he thinks is its socially nihilistic tendencies. Liberalism, with its emphasis on the individual and freedom, has a propensity for the removal of standards that place limits on human thought and action. The result is *moral relativism*, which may remove constraints but leaves people without direction. The modern focus on prosperity and comfort and the social stability necessary for their enjoyment is also troubling to Strauss. These goals are relatively modest when contrasted with the more demanding expectations of premodern theorists. Modernity sets a low bar for human achievement. People are not asked to excel as human beings.

The American philosopher Allan Bloom (1930–1992) adopts Strauss's basic perspective in his influential critique of American higher education in the *Closing of the American Mind* (1987). Bloom claims that the modern political values of liberty and equality have, in part, produced a culture dominated by a complacent and indulgent relativism. Liberty is interpreted simply as freedom from restraint and equality refers to the removal of all hierarchies. The result of no restraints or hierarchy is a lack of standards. The closing of the American mind is a closing to the possibility of truth. Thus, our personal and collective lives do not have real meaning, nor is real meaning possible.

Conservatism is concerned with conserving tradition and custom. There are, however, a variety of conservative political theories. Traditions and customs differ from society to society. Thus, conservative theories often focus on different social contexts and practices as opposed to abstract models of politics. Conservatives may also disagree about the aspects of tradition and custom that should be preserved. Even though conservatism favors conservation, its purpose is not to ossify convention per se but to preserve what is worthwhile. Conservative minds may differ on what is worthwhile and what should be reformed.

The conservative emphasis on tradition reflects both its view of human nature and the threat it perceives from the liberal tradition. Conservatives believe people require the guidance of tradition and support of community due to the inherent shortcomings of humans. Otherwise, the individual will be overwhelmed by the unfamiliar and susceptible to exploitation. Conservatism is also a reactionary theory of politics. Conservative theories often involve critiques of the liberal tradition as dangerous due to its sanguine view of human nature and reliance on abstract reason. They claim the liberal prioritization of freedom undermines the tradition and community needed to structure people's lives. Moreover, conservatism's emphasis on abstract concepts such as the atomized individual, natural rights, and social contract create abstract models of politics that privilege reason over reality. Political institutions and policies must be shaped to fit each society's unique history and traditions.

Perennial Questions and Future Prospects

Political theory, as a normative enterprise, inevitably raises perennial questions. The issue of human nature and what can be realistically expected from people is an ongoing debate. What values are important and how they should be prioritized can never be settled in any objective manner. Thus, modern liberalism, libertarianism, and conservatism will always have to revisit these matters. Proponents will come up with new justifications and rationales. Opponents will continue to raise challenges and critiques.

The modern liberal tradition includes many different theories about how the state should intervene to promote positive liberty and equal opportunity. These differences tend to revolve around three basic questions: (1) What resources does the individual need to enjoy positive freedom? (2) What opportunities ought to be equal? (3) What institutions and policies should be used to satisfy the demands of positive liberty and equal opportunity? These are the perennial questions of modern liberalism. Rawls's *Theory of Justice* (1971) has been particularly influential in shaping contemporary debates about modern liberalism, serving as a foundation or foil for the works of many other important modern liberal theorists. It has also been subject to challenge by any number of liberal critics. Modern liberal perspectives and ideas currently dominate the study of political theory. Thus, modern liberalism will likely remain the central theoretical orientation in the foreseeable future.

Libertarianism has a straightforward goal: minimal government intervention to facilitate negative liberty. Still, the nature of this intervention, even if minimal, is not self-evident. There are going to be perennial questions about the power and responsibilities of the state. Libertarians require that the state protect people from being harmed by others. The criterion of harm, however, is open to interpretation. What constitutes harming another person? Libertarians concerned with a laissez-faire marketplace also have to deal with issues such as the enforcement of contracts, monopoly and competition, and the nature of property rights.

Libertarians play an important role in political theory as effective and consistent advocates of negative liberty in all aspects of social and economic life. However, the number of libertarian political theorists is relatively small. Still, there is reason to be optimistic about the future of libertarianism because of the appeal of negative liberty. There are already nonlibertarians that advocate either laissez-faire capitalism or expansive civil liberties. The task is to convince them that negative liberty and minimal government should be a priority in all areas of human life, not just one particular part.

Conservatism's aversion to abstraction means there is no conservative model of politics that can be used to shape institutions and policies. It consists of basic principles—such as the flawed nature of humans and the respect for tradition—that guide political judgments and practice. These basic principles will always be subject to debate. What is the capacity of humans to be rational and moral? What aspects of tradition should be preserved or reformed? New situations and conditions require new judgments about how to respond and which practices are appropriate.

The future of conservative political theory is tied directly to the liberal tradition for better and worse. Conservatism originally developed as a reaction to liberalism. This reactionary stance remains an important part of

its motivation and substantive focus. The vitality of conservatism continues to come, in part, from its critique of liberalism. Still, there is a question of conservatism's being marginalized by the domination of liberal political ideas. Freedom and equality are powerful values that are now well established in the Western world. Although concepts such as tradition and culture still resonate for more orthodox religious faiths and ethnic groups, liberal ideals are working to erode the conservative aspects of these institutions. How conservatism responds to this liberal challenge will determine whether it remains an important theory of politics or becomes an anachronism.

References and Further Readings

- Berlin, I. (2002). *Liberty*. Oxford, UK: Oxford University Press.
- Bloom, A. (1987). *The closing of the American mind*. New York: Simon & Schuster.
- Burke, E. (1987). *Reflections on the revolution in France*. Indianapolis, IN: Hackett. (Original work published 1790)
- Croly, H. D. (1989). *The promise of American life*. Boston: Northeastern University Press. (Original work published 1909)
- Dewey, J. (1999). *Individualism old and new*. Amherst, NY: Prometheus Books. (Original work published 1930)
- Gray, J. (2000). *Two faces of liberalism*. Oxford, UK: Polity Press.
- Green, T. H. (1986). *Lectures on the principles of political obligation and other writings* (P. Harris & J. Morrow, Eds.). Cambridge, UK: Cambridge University Press.
- Gutmann, A. (1980). *Liberal equality*. New York: Cambridge University Press.
- Hayek, F. A. (1978). *The constitution of liberty*. Chicago: University of Chicago Press. (Original work published 1960)
- Heywood, A. (1999). *Political theory: An introduction* (2nd ed.). New York: St. Martin's Press.
- Hobbes, T. (1994). *Leviathan*. Indianapolis, IN: Hackett. (Original work published 1651)
- Kelly, P. (2005). *Liberalism*. Malden, MA: Polity Press.
- Locke, J. (1983). *A letter concerning toleration*. Indianapolis, IN: Hackett. (Original work published 1689)
- Locke, J. (1988). *Two treatises of government*. Cambridge, UK: Cambridge University Press. (Original work published 1689)
- Maistre, J. de (1959). *On God and society: Essay on the generative principle of political constitutions and other human institutions* (E. Greifer, with L. M. Porter, Trans.). Chicago: Henry Regnery. (Original work published 1809)
- Mill, J. S. (1978). *On liberty*. Indianapolis, IN: Hackett. (Original work published 1859)
- Nisbet, R. A. (1990). *The quest for community: A study in the ethics of order and freedom*. Richmond, CA: ICS Press. (Original work published 1953)
- Nozick, R. (1974). *Anarchy, state, and utopia*. New York: Basic Books.
- Oakeshott, M. (1990). *On human conduct*. Oxford, UK: Clarendon. (Original work published 1975)
- Oakeshott, M. (1991). *Rationalism in politics and other essays*. Indianapolis, IN: Liberty Fund. (Original work published 1962)
- Paine, T. (1982). *Common sense*. New York: Penguin. (Original work published 1776)

- Paine, T. (1992). *The rights of man*. Indianapolis: Hackett. (Original work published 1791 and 1792)
- Plamenatz, J. P. (1938). *Consent, freedom and political obligation*. London: Oxford University Press.
- Pojman, L. P., & Westmoreland, R. (Eds.). (1996). *Equality: Selected readings*. New York: Oxford University Press.
- Rand, A. (2005). *Atlas shrugged*. New York: Dutton. (Original work published 1957)
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Belknap Press.
- Raz, J. (1986). *The morality of freedom*. New York: Oxford University Press.
- Roemer, J. E. (1998). *Equality of opportunity*. Cambridge, MA: Harvard University Press.
- Rothbard, M. N. (2003). *The ethics of liberty*. New York: NYU Press. (Original work published 1982)
- Rothbard, M. N. (2007). *Power and market: Government and the economy*. Auburn, AL: Von Mises Institute. (Original work published 1970)
- Smith, A. (1976). *An inquiry into the nature and causes of the wealth of nations*. Chicago: Chicago University Press. (Original work published 1776)
- Spragens, T. A., Jr. (1975). *Understanding political theory*. New York: St. Martin's Press.
- Strauss, L. (1999). *Natural right and history*. Chicago: University of Chicago Press. (Original work published 1953)
- Von Mises, L. (2007). *Human action: A treatise on economics*. Indianapolis, IN: Liberty Fund. (Original work published 1949)

73

ANARCHISM

LEONARD WILLIAMS

Manchester College

Anarchism, like political science, seems to be both a concept and a practice. As concepts, each has been the subject of a nearly endless series of stipulated definitions. Anarchists and political scientists often appear to be more engaged with defining their approach than with actually doing it. When viewed as practices, though, it is evident that both anarchism and political science have long, convoluted histories in which the past never entirely disappears but is simply rediscovered and revived from time to time. Defining political science, and describing its varieties and folkways, is the subject of this volume. Doing the same for anarchism is the subject of this chapter.

It may seem odd to do so. After all, you will not find anarchist theorists represented in the conventional canon of political philosophers commonly studied in universities. If anarchist thinkers are studied at all, they are addressed largely in courses on political ideologies or occasionally in courses on feminist theory. As carriers of an ideology, either they function as prescient critics of totalitarianism or they appear as unrealistic utopians. At best, anarchism emerges as the solution to an ideological puzzle—one that turns the linear continuum of the left-right spectrum into more of a Möbius strip. At first glance, though, anarchists seem to have little to offer to the academic student of politics. You have to be rather quirky, if not a little defiant, to show any serious interest in their political thought and practice.

This marginalization of anarchists and anarchism in the academy parallels their marginalization and even demonization that has long been present in popular culture. The 19th-century image of the unkempt, bearded, bomb-throwing anarchist is never too far from the popular mind. Anarchists usually preface discussions of their ideas with a caveat that they are not necessarily practitioners of assassination, sabotage, or mindless violence. Even today, whenever shop windows are broken or scuffles with police occur during protest demonstrations, the media are quick to attribute the violence to frightening bands of anarchists such as the Black Bloc.

Little of this is news to any adherent or student of anarchism. Whether as ideological doctrine or revolutionary practice, anarchism has long gone in and out of fashion; its ideas have arisen whenever and wherever revolutionary movements have occurred. Indeed, the classical period of anarchist theory and practice has long been associated with the working-class movements of the 19th century and the antifascist struggles of the Spanish Civil War. In the many decades since, it has been associated with the global student and antiwar movements of the 1960s, with new social movements and indigenous political movements a decade or two later, and more recently, with the antiglobalization movement that came to prominence with the “Battle of Seattle” (Day, 2005). Indications of anarchism’s vicissitudes can be found in any number of essay collections that

attempt to capture the range and spirit of a “contemporary” anarchism (Ehrlich, 1996; Purkis & Bowen, 2004b).

At present, anarchism has experienced yet another rebirth, primarily among two constituencies: a younger generation of radical political activists coming of age in the 1990s and a number of academic thinkers steeped in poststructuralist philosophical perspectives. This renewed appreciation for anarchist theory and practice reveals what Daniel Guérin (1970) observed long ago: “Anarchism can be described first and foremost as a visceral revolt” (p. 13). Because that spirit of revolt recurs from time to time, anarchism must be taken seriously by students of political science (Gordon, 2007; Williams, 2007). As a result, the aim of this essay is threefold: (1) to introduce some of the classic thinkers associated with anarchism, (2) to provide a glimpse into the diverse schools of thought represented in today’s anarchist world, and (3) to sketch key features of an anarchist worldview.

Classic Thinkers

Although its origins can be traced to Enlightenment thinkers (Jean-Jacques Rousseau and William Godwin, for example) or even to Taoist thought, anarchism has been associated primarily with 19th century political and social movements. In this section, we will summarize the central ideas of some key thinkers in the anarchist pantheon.

Pierre-Joseph Proudhon (1809–1865) was the first thinker to proclaim proudly his identity as an anarchist. Famous for declaring (in reference to legal entitlement, not simple possession) that “property is theft,” Proudhon consistently promoted values such as spontaneous organization and free association. The actions of government, any government, violate what he saw as a core principle—the only valid contract is one that has been freely entered into by autonomous individuals. In contrast, Proudhon preferred a society in which people’s interactions were marked by mutuality, cooperation, and reciprocity. As he noted in *What Is Property?* (1840), “Politics is the science of liberty: man’s government of his fellow-man, no matter the name under which it lurks, is oppression: society’s highest perfection lies in the marriage of order and anarchy” (quoted in Guérin, 2005, p. 54). Such a society, built largely on localist or federalist principles, would emerge spontaneously from the overly regulated one we currently inhabit. The abolition of the state, and the subsequent transformation of law into a vehicle for justice, would come only when human beings freely persuade and convince each other of the need to respect the commands of justice.

Where Proudhon sought to articulate the first principles of social organization, Mikhail Bakunin (1814–1876) spent his energies trying to foment a social revolution on behalf of individual freedom. Ever intent on practicing the revolutionary arts, he was often keen on building new

cadres, and he frequently quarreled with other radical thinkers and activists. Bakunin’s vision of the ideal society was communal and collectivist, one built on abolishing the state and recognizing the claims of common ownership and social labor. Bakunin’s own revolutionary version of class struggle anarchism put him in opposition to both authoritarian socialism and bourgeois democracy. Critical of political power per se, he denigrated any society divided into exploiter and exploited, ruler and ruled, master and slave. The proletarian class rule envisioned by Karl Marx and others, Bakunin asserted in *Statism and Anarchy* (1873), would not bring liberation, but would instead reestablish state authority. “Whoever says State necessarily says domination, and, consequently, slavery: a State without slavery, open or concealed, is inconceivable: that is why we are enemies of the State” (quoted in Guérin, 2005, p. 195). Only a thoroughgoing social revolution could wipe away the state and institute a more natural, decentralized, and equal society. Likely to be bloody and violent, as any war would be, the revolution would nevertheless be a spontaneous, mass one—even though it may need some guidance from a cadre of determined professionals.

Peter Kropotkin (1842–1921) was among a group of anarchists who sought to give anarchism a solid intellectual infrastructure and notably contributed an article on it to the 11th edition of the *Encyclopaedia Britannica*. In that piece, he defined anarchism as the “principle or theory of life and conduct under which society is conceived without government—harmony in such a society being obtained, not by submission to law, or by obedience to any authority, but by free agreements concluded between the various groups” (Kropotkin, 1910, para. 1). Kropotkin’s arguments for anarchism were influenced substantially by scientific discoveries in such domains as biological evolution and human anthropology. Drawing on his own research, he asserted that all species exhibit two natural tendencies: one (rooted in the self-assertion or self-preservation of an individual organism) toward competition and survival of the fittest and another (enabling the survival of a species in an environment) toward harmony and mutual aid. The natural tendency toward mutual aid, long neglected by most intellectuals, not only made species survival possible but also was the prerequisite for any sort of social or moral progress. Once the evolutionary role of mutual aid was understood, people would see that it is only by renouncing capitalism, the state, and violence that we could create an egalitarian communist, radically democratic, and pacifist society.

Noted for her tireless advocacy of workers’ and women’s rights, Emma Goldman (1869–1940) remains today an inspirational figure for organizers and feminists alike. Eloquenty urging the forces of labor to unite in their struggles against capital, she saw in anarchist communism a philosophy that would put an end to the conflict between the individual and the society. In her famous 1910 essay

“Anarchism: What It Really Stands For,” Goldman (1969) showed how anarchism represented the only rational alternative to the pernicious influence of religion, property, and the state. Goldman just as eloquently challenged the many obstacles to women’s emancipation. Linking patriarchy and capitalism, she saw in marriage and other forms of trafficking in women little more than the effects of private property and the state. Finding suffrage to be no more than symbolic politics, she pressed strongly for the authentic liberation of women that would come only with sexual liberty and birth control. One of the first to recognize that the personal is indeed political, Goldman acknowledged the “interdependence of collective social transformation and the inner psychological, mental and spiritual liberation of individuals” (Hewitt, 2002, p. 173).

Max Stirner (1806–1856) was the pseudonym adopted by Johann Kaspar Schmidt, a German philosopher and teacher. In *The Ego and His Own* (1845), Stirner advocated a consistent egoism and individualism that stressed the importance of pursuing one’s own welfare. To a constrictive society, which relies on the compulsory authority of the state, Stirner opposed a liberatory association that rests on the free activity of individuals. Where the former demands loyalty, and compels it through law, the latter promotes unbounded individual autonomy. Stirner’s emphasis on radical autonomy extended to his concept of action; neither political nor social, action always takes place on the individual’s own terms—one unites, and separates, freely and frequently. Forsaking revolution, the overthrow of the existing order, Stirner instead promoted a spirit of revolt:

The Revolution has new institutions as its objective. Revolt induces us to no longer let ourselves be governed, but rather to shift for ourselves. Revolt does not look to the “institutions” to come for any wonders. It is a fight against what already exists. (quoted in Guérin, 2005, p. 29)

The very essence of life for Stirner was to freely create and re-create one’s individuality and subjectivity. Only when the stultifying nature of power is exposed, and then undermined, can authentic life emerge.

Contemporary Variants

Among the theorists and activists writing about anarchy today, one can find any number of strains of thought. To be sure, some strains have undergone greater development or have been given more attention than others. Before addressing the ideas of some tendencies found in contemporary anarchism, though, let us first highlight the ideas of perhaps the most well-known theorist who helped keep the anarchist tradition alive in the United States.

Best known, politically, for his incisive critiques of U.S. foreign policy and the mass media, Noam Chomsky has

often been cited as an influential anarchist thinker. Even so, he usually identified himself as little more than a “derivative fellow traveler” (Chomsky, 2005, p. 135). Over time, as anarchist theory and practice revived, Chomsky (2005) gradually warmed to the label—proclaiming that his “personal visions are fairly traditional anarchist ones, with origins in the Enlightenment and classical liberalism” (p. 191). Viewing anarchism as a libertarian form of socialism, opposed to hierarchy and domination, he remained ever focused on the liberatory potential found in advanced industrial and technological societies.

Although many of today’s anarchists regard him highly, Chomsky has often been criticized by fellow anarchists—either for advancing dated notions of workers’ control and self-management or for assuming that the goal is to create a highly industrial, democratic society. Indeed, his work sometimes appears so rationalist and pragmatic that he even suggests that anarchists may need to defend, rather than simply attack, certain state institutions—while nevertheless seeking to democratize them (Chomsky, 2005). In many ways, Chomsky’s political writings focus less on advancing the cause of anarchy than on scrutinizing the presumptions of contemporary political decisions and discourse. He is better at taking arguments apart, showing their contradictions, than he is at discussing questions of value or outlining the features of a new society. Focused as he is on critiquing what governmental leaders do and say, rationalism seems to be at the center of his activity—even when Chomsky states his predilection and preference for anarchism.

Social Ecology

Like Chomsky, the late Murray Bookchin (1971, 1995) has been frequently cited as a leading and influential anarchist theorist. Indeed, he was one of the first thinkers to link environmental and political concerns and to show the interconnections between ecology and anarchism. Labeling his approach “social ecology,” Bookchin saw each domain as marked by participatory freedom, ever-increasing differentiation, mutuality and complexity, and unity in diversity (Biehl, 1997). To embrace social ecology is to denounce hierarchy in the name of creative freedom and enriching diversity; it is to favor renewable energy and human-scale technology, along with decentralized economic and political structures.

Bookchin’s approach to anarchism emphasizes not only a generalized respect for the environment but also a deliberative, Aristotelian conception of politics. Indeed, his preference for rational discourse and radical democracy brought Bookchin considerable criticism from some anarchists who view the movement not merely as antistate but also as broadly antipolitical. Bob Black (1997) is one who voiced this complaint, finding no better example of leftist anarchy’s sterility than in Bookchin’s preference for local government and direct democracy and his adherence to

rationalist ideology and politics as usual. Many young anarchists have thus sought to move anarchism beyond its old leftist affiliations and Bookchin's abstract theorizing and system-building.

Sharing their concerns with rooting out domination, Bookchin initially declared anarchism to be "a libidinal movement of humanity against coercion in any form, reaching back in time to the very emergence of propertied society, class rule, and the state" (Biehl, 1997, pp. 144–145). However, Bookchin later turned away from such global negativity; rather than give free rein to political libido, for example, he issued a reminder that anarchism must be conceived as an organized movement committed to "four basic tenets: a confederation of decentralized municipalities; an unwavering opposition to statism; a belief in direct democracy; and a vision of a libertarian communist society" (Biehl, 1997, p. 170). Bookchin's rationalism could also be seen in two other tenets of social ecology: support for locally organized and human-scale technology and a belief that the course of evolution is directed toward greater complexity.

Primitivism

A decade ago, a profile of activists opposed to gentrification in the Pacific Northwest highlighted John Zerzan as "a leading advocate of primitivism, which goes far beyond matters of how the state is or isn't constructed, considering technology and most of what we consider civilization to be deeply pathological and needing to be eliminated" (Parrish, 1999, para. 6). As ecological and technological issues came to the forefront, it did not take long for anarchism to become nearly identical with primitivism. Primitivists note that the ills we face today—hierarchy, domination, physical and mental illnesses, violence, and ecological destruction—are due not simply to modernity. They are in fact the necessary consequences of civilization itself: "life before domestication/agriculture was in fact largely one of leisure, intimacy with nature, sensual wisdom, sexual equality, and health. This was our human nature, for a couple of million years, prior to enslavement by priests, kings, and bosses" (Zerzan, 1998, p. 16).

Each of our systems—technology, domestication and agriculture, the division of labor, urbanization, and language—places human beings against nature and embeds us in relations of conformity and obedience. Emerging from the bonds of this "megamachine" requires joining our insights into hunter-gatherer societies with traditional anarchist advocacy of revolution. This approach is not without its limitations, for as Michael Albert (2006) has noted, primitivists err in attributing "the problem they pose not to mutable social structures and institutions which impose the bad features on the technologies and the bad technologies on us, but to the entire category of technology per se." Regardless of the merits of this position, primitivists have gained influence because they recognized the

value of the propaganda of the deed—doing something, anything, to oppose the established order. In their eyes, eschewing violence for either moral or tactical reasons does nothing but further entrench the very ills that anarchists oppose. For Zerzan (n.d.), the only "truly humanitarian and pacific impulse is that which is committed to relentlessly destroying the malignant dynamic known as civilization, including its roots."

Ontological Anarchism

In November 1999 in Seattle, a diverse collection of activists coalesced to protest the World Trade Organization and, without hierarchy or bureaucracy, "organized a wide range of activities, including marches, human blockade chains, banner displays, street props and pavement theatre" (Sheehan, 2003, p. 8). Ever since, such forms of protest have become central to the frame of emancipatory politics. In this context, one of the more intriguing thinkers is Hakim Bey (the pseudonym of Peter Lamborn Wilson), whose ideas about anarchism and culture have encouraged activists to embrace an expressive, artistic form of rebellion—an "ontological anarchism."

Setting aside the traditions of leftist revolution, the ontological anarchist seeks to liberate the imagination through spontaneous acts of *poetic terrorism* and *art sabotage*—acts designed to shock people out of their complacency and alter their consciousness. Following lines of thought initially developed by the Situationist International (Knabb, 2006) and representing anarchism as a sort of lived poetry, ontological anarchists have the goal of using whatever tools are available to disrupt the routines of everyday life. As Bey (2003) suggests, we must "murder the IDEA—blow up the monument *inside us*" in order to stimulate a shift in the balance of power (p. 33). At the heart of these efforts lies the Temporary Autonomous Zone (TAZ) as a compelling, nonauthoritarian approach to social change. The TAZ is conceived as an "uprising which does not engage directly with the State, a guerilla operation which liberates an area (of land, of time, of imagination) and then dissolves itself to re-form elsewhere/elsewhen, *before* the State can crush it" (Bey, 2003, p. 99). The TAZ permits momentary acts of rebellion in aspects of social life that the State has yet to reach; it is less about creating the Revolution and more about fostering the excitement of continual insurrection. For some anarchists, the problem with this approach is that it seems more likely to generate acts of personal resistance and chic rebellion than to develop any useful analysis of society or corresponding institutional change. Bookchin (1999) is one of these critics; he criticizes the TAZ perspective as follows:

[The capitalist system cannot] be overthrown by the creation of Temporary Autonomous Zones, or by "closing" down a government or commercial center for a few hours or even a day, or by routine tussles with the police, or by having a street festival with black flags draped from lampposts. (p. 241)

Not only does Bey's notion of the TAZ fail to challenge authority directly; his partiality for nomadic rebellion privileges a vanguard of liberal moral agents who emphasize their autonomy and intermittent social obligations. Partly in response to such criticisms, Bey has since advanced other conceptions of political activity—whether creating vibrant experiences of communion and community (“immediatism”) or promoting political struggles against power and personal struggles for self-control (“jihad”). Whatever form of action he might recommend, Bey always draws on the traditional anarchist spirit of revolt. Ever opposed to the Revolution and the Spectacle, Bey advocates for “psychic nomadism” and aesthetico-political autonomy.

Anarchist Feminism

The ties between anarchism and feminism are longstanding ones. As activists on issues related to family, sexuality, and work, the anarchist women of the 19th century did then what many anarchists do today: argue for individual autonomy and economic independence, engage in political action through affinity groups, and develop alternative social and cultural institutions. The most important tie between anarchism and feminism, though, is that both have developed an extensive critique of power relationships. L. Susan Brown (1996) argued that even though many feminists could criticize power, it was still “possible and not inconsistent for a feminist to embrace the use of power and advocate domination without relinquishing the right to be a feminist” (p. 151). In making this claim, Brown was taking issue with Peggy Kornegger's (1996) view that feminists had long been “unconscious anarchists” and that the “radical feminist perspective is almost pure anarchism” (p. 159). At times, it seemed as though these thinkers were disputing whether feminism or anarchism was the more universal or inclusive radicalism.

Regardless of whether one believes that anarchism encompasses feminism or that feminism subsumes anarchism, both perspectives acknowledge the role played by *intersectionality* (multiple, interlocking systems—e.g., sexism, racism, capitalism, and heterosexism) in structuring social and political domination. With no privileged perspective available, both feminism and anarchism have come to the conclusion that “the focus of analysis (and of resistance) need not be on any *single* relationship of domination and subordination . . . , but rather on *relations of domination and subordination as such*” (Ackelsberg, 1997, p. 164). In the face of a complex institutional and ideological web that maintains domination, the only hope is a long process of resistance struggles, rooted in difference and partiality, in transgressing borders. The path to liberation thus appears as a process of “hollowing out” the system “through the formation of mental and physical (concrete) alternatives to the way things are” (Kornegger, 1996, p. 164).

Postanarchism

In developing an antiauthoritarian politics, anarchist thought today has been shaped both by new social movements in advanced industrial countries and by poststructuralist social and political theory. By the late 1990s, the political thought and actions stimulated by these developments had produced a paradigm shift within the anarchist tradition (Black, 1997; Purkis & Bowen, 2004a). One notable facet of this paradigm shift was the emergence of a “postanarchist” tendency that emphasized “a rejection of essentialism, a preference for randomness, fluidity, hybridity and a repudiation of vanguard tactics, which includes a critique of occidental assumptions in the framing of anarchism” (Franks, 2007, p. 128). Postanarchism might best be seen as an updated version or modification of anarchism rather than a wholesale rejection of its traditional concerns, insofar as several of the relatively “new” poststructuralist themes (e.g., ideas concerning the sources of oppression, generative power, and positive freedom) were actually first expressed by any number of “classical” anarchist thinkers.

The critique and unmasking of pervasive instances of power has certainly been common to both poststructuralism and anarchism. Yet the very antiessentialism that makes poststructuralist thought so appealing carries within it a “theoretical impasse: if there is no uncontaminated point of departure from which power can be criticized or condemned, if there is no essential limit to the power one is resisting, then surely there can be no resistance against it” (Newman, 2001, p. 5). One problem for anarchism, whether imbued with poststructuralist insights or not, is to find some means of overcoming that impasse. The challenge is to find a point of leverage that can overthrow existing structures of domination without also reviving those same structures—or worse, creating new and more insidious ones.

In this context, today's anarchists seem to have little to offer other than “an *affinity for affinity*, that is, for non-universalizing, non-hierarchical, non-coercive relationships based and [*sic*] mutual aid and shared ethical commitments” (Day, 2005, p. 9). More traditional, programmatic anarchists find such micropolitics hedonistic at worst and unhelpful at best. Because social revolution requires political efforts more serious than a dinner party, more long term than a demonstration, the paradox still remains. Given poststructuralist analyses of productive power (which engenders and occurs within a multiplicity of practices), the rejoinder is that political action can no longer proceed on the basis set forth in old narratives of cataclysmic revolution. Instead, radicals must operate within a micropolitics recognizing the local and contingent nature of political life and calling “for social, personal, and political experimentation, the expansion of situated freedom, the release of subjected discourses and genres, and the limitation and reorientation of the role of the intellectual” (May, 1994, p. 112).

Common Ideas

In the face of such diverse perspectives, the first principles of an ideology such as anarchism appear to be an elusive quarry. Even its defenders regard anarchism as more of an evolving tradition—a set of overlapping and sometimes competing traditions or aspects—than a general theory or a coherent ideology. Because it transforms itself to fit time, place, and circumstance, any effort to set forth a contemporary platform for anarchists seems likely doomed to incompleteness, if not failure. Indeed, sometimes it seems that the only thing that is constant about anarchism is its inconstancy, its ability to transform itself. Nevertheless, despite the absence of any universal anarchist credo, enough family resemblances among the various bodies of anarchist thought exist to make it possible for us to talk about anarchism as a discourse or even an ideology.

Philosophical Anarchism

One element of that discourse is philosophical anarchism, which proceeds from a fundamental opposition to the existence of the state and the authority relations that the state codifies, legitimates, or represents. The aim of this sort of anarchism thus seems to focus on providing a foundation for anarchism's central claims about human nature or the state. For example, its roots can be traced to Robert Paul Wolff (1998), who drew on the Kantian tradition to analyze what he regarded as the deep conflict between authority and autonomy. Because people are radically free, wholly autonomous moral beings, there can be no obligation to obey the state or, indeed, any authority. "Anarchism is the only political doctrine consistent with the virtue of autonomy" (Wolff, 1998, p. 18). Another example can be seen in the work of Alan Carter (2000), who argued that anarchism consisted of a normative preference for political equality and an empirical belief that the state cannot provide that equality. In his view, anarchism developed an understanding of economic and political relations (the state-primacy theory, in which state power selects relations of production in its own interests) that better accounts for the history of the state than does analytical Marxism. More recently, Crispin Sartwell's (2008) contribution, beginning from the premise that coercion is evil, proceeds to demonstrate that all arguments for the legitimacy of state power (whether contractarian, utilitarian, or justicial) fail to persuade or convince.

Still other philosophers have explored anarchism, not from traditional philosophical roots and concerns, but from poststructuralist ones (May, 1994; Newman, 2001). These philosophical anarchists draw their inspiration and approaches from the antiessentialism and nonfoundational approaches of such thinkers as Gilles Deleuze, Jacques Derrida, Michel Foucault, and Jacques Lacan (see Newman, 2001). Their recognition that power is deeply embodied in networks of discourses and practices very much comports with the following fact:

The bulk of ongoing anarchist praxis and discourse takes place on the micro level of face to face collectives and affinity groups, and the meso level of the local milieu or (mini) network of anarchists in a particular locale, such as a town or city. (Gordon, 2007, p. 33)

Indeed, what helped make anarchism appealing to poststructuralist thinkers is that it represents "an ethical critique of authority—almost an ethical duty to question and resist domination in all its forms" (Newman, 2001, p. 166). Recognition of that duty encourages Simon Critchley (2007) to conceive of anarchism as a theory and practice of infinite responsibility rather than a philosophy of radical or unlimited freedom. The responsibility is rooted in an ethical demand for political actors to resist injustice. In accepting and meeting that demand, anarchists challenge the state and question authority—whenever and wherever they seek to totalize social relations or dominate political practices. For Critchley (2007), "The great virtue of contemporary anarchist practice is its spectacular, creative and imaginative disturbance of the state" (p. 123). What anarchism brings to politics, in other words, is a pervasive and compelling spirit of revolt.

Anarchist Theory and Practice

Anarchism thus appears to be more an evolving tradition of theory and practice than a coherent doctrine of ideological principles. Constituted by an expanding set of overlapping and competing traditions, anarchism exists more as a practical creed than as a formal ideology. Even so, the ideological core of anarchism, "the one thing without which it is not anarchism, is the negation of authority over anyone by anyone" (Walter, 2002, p. 32). The necessary corollary to this bedrock commitment is anarchists' assertion that humans can be trusted to pursue their own goals, can indeed live in peace and harmony with nature and with other human beings.

Another noteworthy element of anarchism is that it can no longer (if it ever really could) be regarded as a singular, let alone monolithic, movement. Today, especially, anarchism appears as a plural, a movement of movements. Because oppressive coercion comes from multiple, interconnected sources, the theories and practices of liberation must necessarily be multifaceted and open-ended as well. As a result, anarchism has become more of a synthetic ideology than it was in the classical period. By expanding on anarchism's traditional focus on antiauthoritarianism, contemporary anarchists seek a multifaceted goal:

to highlight not only the state but also gender relations, and not only the economy but also cultural relations and ecology, sexuality, and freedom in every form it can be sought, and each not only through the sole prism of authority relations, but also informed by richer and more diverse concepts. (Grubacic, 2006, para. 19)

Finally, anarchism remains very much a tactical theory—a theory of social and political practice. Propaganda of the deed, as opposed to the development of a “scientific socialism,” was a central preoccupation for classical anarchists; among anarchists in the 21st century, that preoccupation has not dissipated. Having largely adopted (consciously or not) poststructuralist perspectives, anarchists believe that doing what one can, wherever one can, however one can, provides the only prospect of making any headway in the battle against the machine. At present, then, anarchist practice seems to focus on building a set of institutions, resources, skills, and experiences that delegitimize authority and induce a change in perspective, all the while insisting that there *is* an alternative to the present order. Thus,

[Anarchism] is not opposed to organization. It is about creating new forms of organization. It is not lacking in ideology. Those new forms of organization *are* its ideology. It is about creating and enacting horizontal networks instead of top down structures like states, parties or corporations; networks based on principles of decentralized, non hierarchical consensus democracy. (Graeber, 2002, p. 70)

Conclusion

This review of anarchist theory and practice, from its classical period to the present, acknowledges that the anarchist tradition has many facets. It is also a tradition of political thought and action that has undergone many a rise and fall. One motive for including a discussion of anarchism in this work on 21st-century political science is that a new generation of activists and thinkers has produced yet another incarnation of the ideology.

Ultimately, what makes anarchism worthy of serious study is that it represents a sort of limit case for political science. As a political theory, anarchism challenges the premises of the subject of our discipline. It calls into question the very existence of the state and regards conventional political life as a dead end. As a political practice, anarchism represents the seemingly endless search for a means to an end. Its goal is to find some way of undermining people's willingness to buy into the established order, to find some leverage capable of transforming social life. In both cases, the ideological complexity of anarchism seems to translate into a simple libertarian core—a fundamental opposition to authority and a pervasive spirit of revolt.

To be sure, the prospects for ultimate liberation appear to be dim. There may always be some form of authority to combat, some line of conflict or other to overcome. As soon as one perceived problem or division appears to be resolved, another will take its place. As a result, many concur with Stuart White's (2007) conclusion that “the practical role of the anarchist is not to build this unattainable dream, but to push the messy complexity of society in a more anarchist direction” (p. 24). For that pursuit, we can always draw on a rich tradition of ideas that can be

readily counterposed to any orthodoxy—as the current revival of anarchism demonstrates.

References and Further Readings

- Ackelsberg, M. (1997). Rethinking anarchism/rethinking power: A contemporary feminist perspective. In M. L. Shanley & U. Narayan (Eds.), *Reconstructing political theory: Feminist perspectives* (pp. 158–177). University Park, PA: Pennsylvania State University Press.
- Albert, M. (2006). *Anarchism Zerzan? ZNet*. Retrieved May 24, 2010, from http://www.zcommunications.org/anarchism_zerzan_by_michael_albert
- Bey, H. (2003). *T. A. Z.: The temporary autonomous zone, onto logical anarchy, poetic terrorism* (2nd ed.). Brooklyn, NY: Autonomedia.
- Biehl, J. (Ed.). (1997). *The Murray Bookchin reader*. London: Cassell.
- Black, B. (1997). *Anarchy after leftism*. Columbia, MO: Columbia Alternative Library.
- Bookchin, M. (1971). *Post scarcity anarchism*. Berkeley, CA: Ramparts Press.
- Bookchin, M. (1995). *Social anarchism or lifestyle anarchism: An unbridgeable chasm*. Oakland, CA: AK Press.
- Bookchin, M. (1999). *Anarchism, Marxism, and the future of the left: Interviews and essays, 1993–1998*. Edinburgh, UK: AK Press.
- Brown, L. S. (1996). Beyond feminism: Anarchism and human freedom. In H. J. Ehrlich (Ed.), *Reinventing anarchy, again* (pp. 149–155). Edinburgh, UK: AK Press.
- Carter, A. (2000). Analytical anarchism: Some conceptual foundations. *Political Theory*, 28(2), 230–253.
- Chomsky, N. (2005). *Chomsky on anarchism*. Edinburgh, UK: AK Press.
- Critchley, S. (2007). *Infinitely demanding: Ethics of commitment, politics of resistance*. London: Verso.
- Day, R. J. F. (2005). *Gramsci is dead: Anarchist currents in the newest social movements*. London: Pluto Press.
- Ehrlich, H. J. (Ed.). (1996). *Reinventing anarchy, again*. Edinburgh, UK: AK Press.
- Franks, B. (2007). Postanarchism: A critical assessment. *Journal of Political Ideologies*, 12(2), 127–145.
- Goldman, E. (1969). *Anarchism and other essays*. New York: Dover.
- Gordon, U. (2007). Anarchism reloaded. *Journal of Political Ideologies*, 12(1), 29–48.
- Graeber, D. (2002). The new anarchists. *New Left Review*, 13, 61–73.
- Grubacic, A. (2006, May 11). *Power and revolution: The anarchist century*. ZNet Retrieved June 20, 2008, from <http://www.zcommunications.org/znet/viewArticle/3899>
- Guérin, D. (1970). *Anarchism: From theory to practice*. New York: Monthly Review Press.
- Guérin, D. (Ed.). (2005). *No gods, no masters: An anthology of anarchism*. Oakland, CA: AK Press.
- Hewitt, M. (2002). Emma Goldman: The case for anarcho feminism. In D. Roussopoulos (Ed.), *The anarchist papers* (pp. 167–175). Montréal, QC, Canada: Black Rose Books.
- Knabb, K. (Ed.). (2006). *Situationist international anthology* (Rev. and expanded ed.). Berkeley, CA: Bureau of Public Secrets.

- Kornegger, P. (1996). Anarchism: The feminist connection. In H. J. Ehrlich (Ed.), *Reinventing anarchy, again* (pp. 156–168). Edinburgh, UK: AK Press.
- Kropotkin, P. (1910). "Anarchism," from the *Encyclopaedia Britannica*, 1910. *Anarchy Archives*. Retrieved May 26, 2009, from http://dwardmac.pitzer.edu/anarchist_archives/kropotkin/britanniaanarchy.html
- May, T. (1994). *The political philosophy of poststructuralist anarchism*. University Park, PA: Pennsylvania State University Press.
- Newman, S. (2001). *From Bakunin to Lacan: Anti authoritarianism and the dislocation of power*. Lanham, MD: Lexington Books.
- Parrish, G. (1999, 2–8 September). The new anarchists. *Seattle Weekly*. Retrieved May 23, 2010, from http://www.seattleweekly.com/1999_09_01/news/the_new_anarchists/
- Purkis, J., & Bowen, J. (2004a). Introduction: Why anarchism still matters. In J. Purkis & J. Bowen (Eds.), *Changing anarchism: Anarchist theory and practice in a global age* (pp. 1–19). Manchester, UK: Manchester University Press.
- Purkis, J., & Bowen, J. (Eds.). (2004b). *Changing anarchism: Anarchist theory and practice in a global age*. Manchester, UK: Manchester University Press.
- Sartwell, C. (2008). *Against the state: An introduction to anarchist political theory*. Albany: SUNY Press.
- Schmidt, M., & van der Walt, L. (2009). *Black flame: The revolutionary class politics of anarchism and syndicalism*. Oakland, CA: AK Press.
- Sheehan, S. M. (2003). *Anarchism*. London: Reaktion Books.
- Walter, N. (2002). *About anarchism* (Updated ed.). London: Freedom Press.
- Ward, C. (2004). *Anarchism: A very short introduction*. Oxford, UK: Oxford University Press.
- White, S. (2007). Making anarchism respectable? The social philosophy of Colin Ward. *Journal of Political Ideologies*, 12(1), 11–28.
- Williams, L. (2007). Anarchism revived. *New Political Science*, 29(September), 297–312.
- Wolff, R. P. (1998). *In defense of anarchism*. Berkeley, CA: University of California Press.
- Zerzan, J. (1998). *Future primitive and other essays*. Brooklyn, NY: Autonomedia.
- Zerzan, J. (n.d.). *On the transition: Postscript to "Future Primitive."* Retrieved July 19, 2007, from <http://www.primitivism.com/transition.htm>

74

NATIONALISM

SALIH BİÇAKCI

Isik University

Nationalism is a modern ideology that tries to explain the individual's devotion to the nation-state by neglecting other interests. It has taken many different shapes in various geographies, cultures, histories, and political systems. Even in a particular location, nationalism has transformed from one form to another throughout history. The core of nationalism is *nation*. What constitutes a nation is a question scholars are still trying to clarify by using approaches developed throughout the ages. Not only political science but also other branches of the humanities are trying to understand the concepts of nation and nationalism. The current technological innovations and rapid globalization have added new dimensions to nationalism and its movements. Each day brings a new peculiarity of nationalism.

The various definitions recall the story in which a group of blind men touch an elephant to learn what it is like. Each one touches a different part, but only one part, such as the side or the tusk. They then compare notes on what they felt and learn they are in complete disagreement. All attempts to define nationalism are similar: They come from the perspective of the scholars' disciplines, and like the blind men, each discipline touches only one aspect of nationalism. As a result, a remarkable amount of research has been published regarding nationalism, but theoretical progress has been limited.

The concept of "nation" is historically older than nationalism as a political movement. The English word *nation*

comes from the Latin word *nasci*, which literally means "to be born." The word has gradually taken the meaning of large group of people with a common ancestry. The idea of nation takes shape in conjunction with cultural, political, and psychological factors.

Language, religion, history, literature, folkloric themes (epics, myths, legends), and customs are the elements creating bonds among a group of people that transform a nation. Indeed, there is no consensus among scholars and researchers on the subjective and objective factors for the definition of *nation*. Anthony Smith (2001) distinguishes the objective factors of language, religion, customs, territory, and institutions from the subjective category of attitudes, perceptions, and sentiments. Renan (1882) identified the nation as a form of morality and solidarity that was supported by historical consciousness. On the other hand, Max Weber agrees that the nation is "obviously an ambiguous term" (quoted in Gerth & Wright-Mills, 1948, p. 176). But his way of understanding takes us to the point at which his nation concept becomes a prestige community unified around a myth of common descent. Weber also understands the nation as a political project that "tends to produce a state of its own" (p. 177). On the other hand, Stalin expounded on the nation as a combination of subjective and objective elements. According to Stalin, "A nation is a historically constituted, stable community of people, formed on the basis of a common language, territory, economic life, and psychological make-up manifested in a

common culture” (quoted in Franklin, 1973, p. 57). From a different perspective, Greenfeld (1992) states that “social, political, and cultural in the narrow sense, or ethnic qualities, acquire a great significance in the formation of every specific nationalism” (p. 8).

The idea of a nation as a cultural entity dates back to 18th-century German political thinkers. Johann Gottfried von Herder, a critic, poet, and philosopher, was the first author to mention that each nation had a cultural distinctiveness (Hayes, 1927; White, 2005). He emphasized the importance of language and asserted that certain ideas of an individual in one language could not be understood in another language. He also believed that language constructed one’s worldview (*weltanschauung*). He demonstrated how epics, myths, legends, and folk songs build a spirit that can be named *volksgeist*. Herder preferred to refer to it as the “spirit of nations” (*Geist des volkes*). Herder collected folk songs, which he published in his work *Voices of the People in Their Songs* to underline the value of national culture, collective memories, and traditions for a nation (Herder, 1818). The definition of the German romantic writers was criticized with the claim that cultural commonalities were not as powerful as in agricultural societies. On the contrary, modern researchers underlined the role of the industrial revolution and modernization in the spread of nationalism. Ernest Gellner (1983), a modern philosopher, defined “nationalism as primarily a political principle that holds that the political and the national unit should be congruent” (p. 1). He interpreted the new cultural cohesion as a product of the industrial revolution promulgated by education and the division of labor in industry. He conceptualized the culture of the industrial age as high culture transmitted through education (Gellner, 1983).

German historian Karl Renner (Reifowitz, 2009) added another level to the discussion by demonstrating how historical destiny transformed “passive people” (*passiver Volkheit*) into a group that had become conscious about itself (Renner, 1899, p. 89, quoted in Hobsbawm, 1990, p. 101). Anthony Smith (1983) followed a similar approach by explaining modern nationalism in relation to premodern ethnicities. He claimed that a nation is embedded in the history of its cultural, linguistic, and political values. This type of self-affirmation also led to cultural nationalism with a primary emphasis on cultural distinctiveness. Another German historian, Friedrich Meinecke (1919), clarified the modern state-and-culture relation by identifying *Kulturnation* as a “largely passive cultural community” and the *Staatsnation* as an “active self-determining political nation” (pp. 2–3). He identified the nation as cultural or ethnic affiliation versus the nation as political state. Meinecke referred to the Germans, the Russians, the Irish, the Greeks, and the English as examples of *kultur nation*. From this perspective, since culture cannot be learned, it is not possible to become German by learning the language and adopting the lifestyle and values. You have to be a native German to perceive the culture. This

distinction also implies two enduring ways of understanding the rise of the nation-state.

The first appearance of European nationalism has been a topic of discussion. In 1648, at the end of the Thirty Years’ War, the European powers signed the Treaty of Westphalia in Munster and Osnabruck, ending interference in each other’s domestic politics. The principle rule—*cuius regio, eius religio* (whose realm, his religion)—of the treaty confirmed that the ruler’s faith became the official religion of his state. The states formed on this principle were accepted as preliminary examples of the nation-state in the political science literature (Schulze, 1998). Unlike other researchers who have taken the French revolution as a first example, Greenfeld (1992) argues that “the original modern idea of the nation emerged in sixteenth century England, which was the first nation in the world (and the only one with the possible exception of Holland, for about two hundred years)” (p. 14).

Political Dimension of Nations

The political allegiance, citizenship, and homogeneous population that form the nation are products of the modern age. A nation is a group of people bonded to each other by citizenship under the authority of a political construction that ignores cultural, ethnic, and other loyalties. In this sense, Andrew Heywood (2000) basically understood the nation as a *psycho political* construction. But it also has a historical progress dating back to the French revolution, when the transition from monarchic structures, in which the individuals were subjects of the crown, to the constitutional state, which promoted participatory rule, took place. Jean-Jacques Rousseau conceptualized the participation of the people with the term *general will*. In *Social Contract*, Rousseau (1762/2008) wrote, “Each of us puts his person and all his power in common under the supreme direction of the general will, and, in our corporate capacity, we receive each member as an indivisible part of the whole” (p. 77). In these lines, he explains general will not only as one of the principles of the nation but also as a condition for the formation of the nation-state. Nationalism emerged from the national sentiment created within these nation-states. Mainstream researchers understood nationalism as an output of nation-states. Meinecke created the term *state nation* to describe an entity differing from the nation-state. The concept of state-nation was based on Rousseau’s idea of general will and the nations formed by states. According to Meinecke, the nation-state gradually evolved from an individual culture. As a result, he concluded that states were formed from nations.

Eric Hobsbawm (Hobsbawm & Ranger, 1983) opposed the idea that nations were basically ethnic groups formed throughout history. He asserted that nations were superficially formed by nationalism, and he conceptualized

the condition as an “invented tradition.” He presented his example thusly:

Israeli and Palestinian nationalism or nations must be novel, whatever the historic continuities of Jews or Middle Eastern Muslims, because the very concept of territorial states, of the current standard type in their region, was barely thought of a century ago, and hardly became a serious prospect before the end of World War I. (pp. 13–14)

Benedict Anderson’s research also supports Hobsbawm with his use of the term *imagined communities*. Anderson (1991) stated that “a nation is an imagined political community that is imagined as both inherently limited and sovereign.” He also clarified his approach thusly: “A nation is imagined because the members of even the smallest nation will never know most of their fellow-members, meet them, or even hear of them, yet in the minds of each lives the image of their communion” (pp. 5–6). He claimed that education, political communication, and the mass media played a crucial role in building this imaginary sense. These approaches have also been supported by Marxism, which believes that the concepts of nation and nationalism belong to the bourgeoisie. These concepts were constructed as instruments to exploit (rule) other classes through the creation of loyalty based on the sense of nation, which was more powerful than the binding power of the working classes.

Nations somehow evolved into politics and have thereafter been processed under the rules of politics. As in Meinecke’s definition of political nations, the significance of citizenship is more intensive than that of ethnicity. Cultural heterogeneity is one of the common indicators of these countries. The United States and the United Kingdom are given as the examples of this type of political nation. In this context, Meinecke also differentiates the terms *state nation* and *nation state*. The nation-state refers to the state that was built on the crystallization of an individual culture. However, the state-nation is based on Rousseau’s “general will” and is a nation constructed by the state. The case of the United States fits the state-nation concept. It is hard to build a national identity that depends on the commonality of a shared cultural and historical past because of the multiethnic and multicultural characteristics of the United States. U.S. nationhood formed around the voluntary acceptance of a set of common values, principles, and goals by all citizens. It is possible to use the melting-pot analogy for these types of states. Since state-nations are not composed of one individual culture, they have the challenge of creating an organic unity.

Nationalism and the political nation concepts have generally been understood in the European context. As a result, the nation-state and national identity have peculiar problems in the third world, where two major streams have been followed. First, national identities were built up during their struggles for freedom in national independence wars

against colonial powers. These identities were strongly shaped under the anticolonial characteristics of that period. Second, national identity was shaped by territorial boundaries. These borders were usually inherited from the colonial past. Contemporary maps of the Middle East and Africa provide a clear example of these divisions. These “nations” have a wide range of ethnicities, but few commonalities except their shared colonial past. Therefore, to achieve statehood, “nationhood” had to be built on existing conditions, which rewrite the history, fabricate a national language, and produce a national education system. Nevertheless, the differences in ethnic and political identities generate tensions within the nation and, from time to time, escalate into conflict. The transformation from colonial rule and empires to nation-states affected the nationalism movements of the 20th century.

Theories of nationalism contain the tensions of the definition of a nation, whether it is the old, naturally given sense or a fabricated output of modernity. Theories of nationalism can be grouped into two major categories. The first group is bounded by two opposing ideas: instrumentalism and primordialism. *Primordialism* mainly understands nationalism as a natural process that stems from such givens as the same blood, language, religion, kinship, and common fate. This type of nationalism is also known as *organic nationalism*. Primordial nationalism approaches the natural nations as having been present throughout history as major actors that played a critical role in shaping the modern world.

On the other hand, *instrumentalism* explains the nation as a product of elite manipulation and concludes that nations can be fabricated. According to the instrumentalist view, the concept of a natural nation is the purposive output of the processes of mental production. The instrumentalist also believes that the statesmen and the elite of the bureaucracy construct the nation and nationalism as “a strategic device to be utilized for political gain, as a great manipulating force that can be used in order to corral a population into a desired position” (Ozkirimli, 2000, p. 86). In his book *Theories of Nationalism*, Umut Ozkirimli (2000) summarized the case thusly: “True instrumentalists believe that nationalism comes from the state, not the other way around” (p. 86). Instrumentalism focuses on the functional capabilities of nationalism. In practice, these two theories are mainly implemented more on ethnicity and ethnic identity than on nationalism.

The second major group of nationalism theories, perennialism and modernism, focus on nations and nationalism. *Perennialism* accepts that nationalism is a modern concept but insists that ethnic communities and cultural identities have existed in all periods of history. Different from the primordialist, the perennialist claims that nations or ethnicities are not natural givens but historical, social, and cultural phenomena. Perennialists view modern nations as updated versions of ethnic communities. However, the *modernist* approach believes that the nation and nationalism emerged

in the modern period as a result of structural changes in societies during the transition to modernity. The modernists also emphasize the change of social institutions and how that change affects society in terms of nationalism.

The Political Presentation of Nationalism

The application of theories on the ground generates varieties of nationalism in political life. Nationalism seems to have been one of the most progressive and driving forces of political life in the 20th century. However, the characteristics of nationalism are generally shaped in the context of and according to the political ideas attached to it. Nationalism could be progressive, liberating, reactionary, authoritarian, conservative, democratic, oppressive, left wing, or right wing. In a colony, it emerges as anticolonial nationalism and promotes the liberation of the people. The major political presentations of nationalism could be grouped as liberal nationalism, conservative nationalism, expansionist nationalism, and anticolonial nationalism.

Liberal nationalism, also known as *civic* or *civil nation alism*, is different from the other types of political presentations of nationalism in that liberal nationalism is characteristically nonxenophobic. The main assumption of liberal nationalism is that human beings naturally divide into nations that possess a separate and unique identity. Liberal nationalism supports every nation's right to self-determination and freedom. The builder of Italian unification, Giuseppe Mazzini, defined the harmony of nationhood in his work as "the idea of a sisterhood of nations" (Snyder & Montgomery, 2003, p. 322). Liberal nationalism is opposed to oppressive and autocratic multinational empires. For example, Mazzini wished not only to unify Italy but also to throw out autocratic Austrian rule. After World War I, Woodrow Wilson's principles also emphasized the character of liberal nationalism by respecting the rights of nations. The progressive aspect of liberal nationalism mainly appears in its promotion of unity, not rivalry, among nations. It also fosters an environment of peace at the state level, which also influences the international system. In this way, liberal nationalism brings universalism with its scope over nations. Indeed, this approach stemmed from liberals' fear that otherwise the international system would degenerate, causing wars and conflicts. The criticism of liberal nationalism mainly considers the approach romantic and unsophisticated because it concentrates only on the progressive side of nationalism while neglecting tribalism, xenophobia, and racism. Other major critics have focused on the political practices of nation-states without considering how these coincide with the linguistic, religious, and ethnic areas within a state's borders.

Conservative nationalism reached its zenith after World War I with the establishment of national states, although it could be dated back to Bismarck's German nation and

Disraeli's one nation concept. Conservative nationalism brings social cohesion and public solidarity derived from patriotism into focus and is not interested in the national self-determination of liberal nationalism. After the establishment of a nation-state, the political elite of that state try to build a nation through the creation of a consistent history and language. Conservative nationalism takes its power from a shared past, building on its values and institutions. Thus nationalism becomes evident with its traditionalism and nostalgia. The acts of protection of conservative nationalism also grow with the growth of its perceived enemies and suspicion. It furthermore boosts intolerance, bigotry, and zealotry. In this type of nationalism, insiders and outsiders are very clearly differentiated. Conservative governments and their elite are capable of aggressively using the military and foreign policy for expansion. Contrary to liberal nationalism, conservative nationalism does not promise a peaceful political system at either the state or the international level.

One of the major political manifestations of nationalism is *expansionist nationalism*. This type has an aggressive character coupled with intentions to extend its territory. Governments that pursue expansionist policies explain their interest in the territories either with historical causes or with claims that the existing territory the nation inhabits is too small or is not able to physically or economically support the nation's population. Expansionist nationalism became visible before World War II with the examples of Germany and Japan. Expansionist nationalism appears with right wing ideologies and emphasizes the importance of the nation over the individual. Another nation or race is defined as a threat or enemy, and this fear is used in building a national identity sustained by a type of negative integration. The sense of the "other" is the main force and motivation to keep "us" together. The image of the other is formed by prejudgments and negative feelings. In the literature, expansionist nationalism is also referred to as *integral nationalism* and was first mentioned by Charles Maurras (Bushman, 1939), a French nationalist. Throughout history, national liberation struggles have led to extreme nationalism. The last step of expansionist nationalism is to define a natural space for the nation, as in the example of Nazi Germany's demand for lebensraum (living space).

Anticolonial nationalism emerged at the end of the struggle against colonial powers. Early appearances of anticolonial nationalism imitated European forms of nationalism but displayed peculiar characteristics. Each instance of anticolonial nationalism was unique and carried a spatial characteristic. Anticolonial nationalism built on the idea of nationhood by the degree of the exploitation and inequality to which the nations that had a colonial past were exposed. Socialism and particularly Marxism–Leninism were embraced in anticolonial nationalism. During the cold war, the peak of nationalism was the rise of Arab nationalism, which was led by Gamal 'Abd al-Nasir

(Dawisha, 2003). He aimed to unite Arabs to fight a common enemy and to reach common goals. Nasir's radio speeches stole the hearts and minds of the Arab people. Arab nationalism reached its peak just before the Arab-Israeli wars, which also empowered the unity of Arab nations. However, the defeat of Arab states in 1967 initiated the dissolution of the movement. Arab nationalism quickly gave birth to a new ideology known as *Ba'athism* (literally, resurrection), which sought the promotion of pan-Arab socialism (Tibi, 1997). Later, the *Ba'ath* movement continued its presence in Syria, Iraq, and Lebanon with the motto of "Unity, Freedom and Socialism." Since the bipolarity of the cold war played a role in regional politics, the Soviets liked the idea of the establishment of an Arabic socialist party and supported its political presence in the region.

The bipolarity of the cold war ended with the dissolution of the Soviet Union. Its former constituent states then became independent, and this development has spurred the debate on nationalism. At least 18 states have come into existence, 14 of them out of the Soviet Union. Under the Soviet Union, political leaders heavily emphasized soviet identity over national identity. History, literature, myths, and other values were constructed on union membership. In the early years of the Soviet Union, new histories were written to emphasize the unity of the soviets (Edgar, 2004). However, after 1991, the newly independent states felt the necessity to build nationhood and debated hotly the objective and subjective factors to choose in order to become a nation. In this discussion, they rewrote their history books and common values up to the present time. Nevertheless, the multiethnic and religious character of these states formed their biggest obstacle in naming their nation. On the other hand, ethnic nationalism was also enhanced in the post-Communist states. Yugoslavia, and how it divided, was the most vivid example of rising ethnic nationalism. The remarkable shifts in the population due to the immigration of a workforce, rapid industrialization, and urbanization expedited the emergence of ethnic conflicts in the former Yugoslavia (Denitch, 1996). In addition to these trends, the transition from authoritarian rule to democracy triggered national exclusivism and ignited conflict between Serbs, Bosnians, and Croats. In 1992, the Serbian militia systematically killed many Bosnians to capture cities and terrorized families, causing them to flee from their homes. This action was labeled *ethnic cleansing* in the literature (Carmichael, 2002). The deployment of 60,000 soldiers of the North Atlantic Treaty Organization established a cease-fire, but the conflict left 200,000 casualties and 2 million refugees.

The destabilization of Yugoslavia and the ensuing power vacuum due to the loss of authority promoted ethnic nationalism on the ground. The U.S.-led intervention into Iraq in 2003 crystallized the ethnicities and religious sects in the region. Iraq became divided between Sunni Arabs, Shi'ite Arabs, Kurds, and Turkomen (Fontana, 2010;

Marr, 2007). The chaotic security environment of the transition period greatly provoked the condition, and the mobilization of the different ethnic groups began. The establishment of a new central authority has lessened the tension among ethnic groups, but it has not totally ended and continues as low-intensity, ongoing conflict. In the 20th and 21st centuries, the conflicts, coupled with the rise of nationalism, could be seen in the examples of Afghanistan, Rwanda, Nigeria, Macedonia, Transnistria in Moldova, and the Caucasus.

After the end of the cold war, the newly independent states displayed extreme hatred toward their colonial past, as well as rising nationalism. These cases of anticolonial nationalism were mostly engaged by, and presented with, Islam. Religious communities and groups were used to initiate anticolonial struggles and to neutralize the comprador bourgeoisie.

To sum up, in the 21st century, the subjective and objective factors of nationalism are rapidly changing with globalization and technological innovations. Indeed, with its Internet communication capabilities and mass media, the digital age has made the world smaller. Now even the untouched spaces on earth have been connected by global information, which squeezes the local culture in order to accommodate itself. Indigenous cultures are forming counterreactionary identities, and micro-scale nationalisms are emerging. In the long run, the number of small nationalities will probably increase. From the regional perspective, these local nationalisms also unify and create more powerful regional nationalist movements, as well. Because borders are changing and new identities are emerging, social scientists of this century are witnessing how the process of nation building commences and continues in various parts of the world. Local languages are more apparent and supported by international organizations. Oral literatures have been published as books, and rituals of culture are turning into traditions. On the other hand, the concept of the nation is changing, especially in the sense of ethnicity. Thanks to the advances in DNA research, several projects now aim to find the genetic sources of various ethnic groups. Today, it is growing easier to follow the traces of a nation's birth and development, a trend that might modify the meaning of nation and nationalism. Social, technological, and economic challenges are jeopardizing the nation-state concept. It is expected that the term will deviate slightly from its original meaning. Gellner (1992) has described the world we are living in as follows:

[a] world in which one style of knowledge, though born of one culture, is being adapted by all of them, with enormous speed and eagerness, and is disrupting many of them, and is totally transforming the milieu in which men [*sic*] live. (p. 78)

In this context, the definition and study of nationalism are also in transition, and one does not expect them to

settle down soon. But nationalism will be the center of various discussions in world politics and will be the core of political science.

References and Further Readings

- Anderson, B. (1991). *Imagined communities*. New York: Verso.
- Armstrong, J. (1982). *Nations before nationalism*. Chapel Hill: University of North Carolina Press.
- Bhabha, H. K. (1990). Introduction. In H. K. Bhabha (Ed.), *The nation and narration* (pp. 1–7). London: Routledge.
- Brubaker, R. (1996). *Nationalism reframed*. New York: Cambridge University Press.
- Buthman, W. C. (1939). *The rise of integral nationalism in France: With special reference to the ideas and activities of Charles Maurras*. New York: Columbia University Press.
- Carmichael, C. (2002). *Ethnic cleansing in the Balkans: Nationalism and the destruction of tradition*. New York: Routledge.
- Chatterjee, P. (1986). *Nationalist thought and the colonial world*. London: Zed Books for the United Nations University.
- Connor, W. (1972). Nation building or nation destroying? *World Politics*, 24, 319–355.
- Dawisha, A. (2003). *Arab nationalism in the twentieth century: From triumph to despair*. Princeton, NJ: Princeton University Press.
- Denitch, B. (1996). *Ethnic nationalism: The tragic death of Yugoslavia*. Minneapolis: University of Minnesota Press.
- Deutsch, K. (1966). *Nationalism and social communication: An inquiry into the foundations of nationality*. New York: Wiley.
- Edgar, A. L. (2004). *Tribal nation: The making of Soviet Turkmenistan*. Princeton, NJ: Princeton University Press.
- Feifowitz, I. (2009). Otto Bauer and Karl Renner on nationalism, ethnicity and Jews. *Journal of Jewish Identities*, 2(2), 1–9.
- Fontana, G. (2010). Creating nations, establishing states: Ethno religious heterogeneity and the British creation of Iraq in 1919–23. *Middle Eastern Studies*, 46(1), 1–16.
- Franklin, B. (1973). Marxism and the natural question. In B. Franklin (Ed.), *The essential Stalin: Major theoretical writings 1905–1952* (pp. 57–61). London: Croom Helm.
- Gellner, E. (1983). *Nations and nationalism*. Ithaca, NY: Cornell University Press.
- Gellner, E. (1992). *Postmodernism, reason and religion*. London: Routledge.
- Gerth, H. H., & Wright Mills, C. (1948). *From Max Weber: Essays in sociology*. London: Routledge & Kegan Paul.
- Greenfeld, L. (1992). *Nationalism: Five roads to modernity*. Cambridge, MA: Harvard University Press.
- Grosby, S. E., & Leoussi, A. S. (Eds.). (2007). *Nationalism and ethnosymbolism: History, culture and ethnicity in the formation of nations*. Edinburgh, UK: Edinburgh University Press.
- Hayes, C. (1927). Contributions of Herder to the doctrine of nationalism. *American Historical Review*, 32(4), 719–736.
- Hayes, C. (1931). *The historical evolution of modern nationalism*. New York: Smith.
- Herbst, J. (2000). *States and power in Africa: Comparative lessons in authority and control*. Princeton, NJ: Princeton University Press.
- Herder, J. G. (1807). *Stimmen der Völker in Liedern* [Voices of the people in songs]. Tübingen, Germany: Hg. Johannes von Müller.
- Heywood, A. (2000). *Key concepts in politics*. New York: Palgrave.
- Hobsbawm, E. (1990). *Nations and nationalism since 1780*. Cambridge, UK: Cambridge University Press.
- Hobsbawm, E., & Ranger, T. (Eds.). (1983). *The invention of tradition*. Cambridge, UK: Cambridge University Press.
- Horowitz, D. (1985). *Ethnic groups in conflict*. Berkeley: University of California Press.
- Kedourie, E. (1971). *Nationalism in Asia and Africa*. London: Weidenfeld and Nicolson.
- Kohn, H. (1929). *A history of nationalism in the East*. New York: Harcourt, Brace.
- Kymlicka, W. (2001). *Politics in the vernacular: Nationalism, multiculturalism and citizenship*. New York: Oxford University Press.
- Marr, P. (2006). Kurds and Arabs, Sunnis and Shi'ites: Can an Iraqi identity be salvaged? In D. Little & D. K. Swearer (Eds.), *Religion and nationalism in Iraq: A comparative perspective* (pp. 63–72). Cambridge, MA: Harvard University Press.
- Meinecke, F. (1919). *Weltbürgertum und Nationalstaat. Studien zur Genesis des Deutschen Nationalstaates* [The middle classes of the world and the national state: Studies on the genesis of the German national states]. Munich, Germany.
- Neuberger, B. (1986). *National self determination in post colonial Africa*. Boulder, CO: Lynne Rienner.
- Ozkirimli, U. (2000). *Theories of nationalism*. New York: Palgrave.
- Reifowitz, I. (2009). Otto Bauer and Karl Renner on nationalism, ethnicity & Jews. *Journal of Jewish Identities*, 2(2), 1–19.
- Renan, E. (1882). *Qu'est ce qu'une nation?* [What is a nation?] (I. M. Snyder, Trans.). Paris: Calmann Levy.
- Renner, K. (1899). *Staat und Nation* [State and nation]. Vienna, Austria.
- Robinson, D. E. (2001). *Black nationalism in American politics and thought*. New York: Cambridge University Press.
- Rousseau, J. J. (2008). *The social contract or principles of political right* (S. M. Soares, Ed., G. D. H. Cole, Trans.). Retrieved May 23, 2010, from <http://www.constitution.org/jjr/socon.htm> (Original work published 1762)
- Schulze, H. (1998). *States, nations and nationalism: From the Middle Ages to the present*. Oxford, UK: Wiley Blackwell.
- Smith, A. D. (1983). *Theories of nationalism*. London: Holmes & Meier.
- Smith, A. D. (1998). *Nationalism and modernism*. New York: Routledge.
- Smith, A. D. (2001). *Nationalism: Theology, ideology, history*. Cambridge: Polity Press & Blackwell Publishing.
- Snyder, J. (2000). *From voting to violence: Democratization and nationalist conflict*. New York: W. W. Norton.
- Snyder, L. L., & Montgomery, J. D. (2003). *The new nationalism*. New Brunswick, NJ: Transaction.
- Tibi, B. (1997). *Arab nationalism: Between Islam and the nation state*. New York: St. Martin's Press.
- White, R. (2005). Herder: On the ethics of nationalism. *Humanitas*, 18(2), 166–181.

75

FASCISM AND NATIONAL SOCIALISM

GERALD L. McCALLISTER

University of North Texas

The Great War, also known as World War I or the “war to end all wars,” brought the concept of total war to the battlefield, unleashing unprecedented destruction and leaving millions of victims in its wake. After such devastation, it might have been reasonable to expect those affected to be pacified by a feeling of war weariness, but instead we saw the rise of a political ideology whose followers advocated perpetual conflict. Fascism is a quasi-religious political ideology that is anticommunist, antiliberal, anticapitalist, anti-intellectual, antipositivist, anti-internationalist, anti-Christian, anticonservative, anti-rationalistic, antiproletarian, antibourgeois, anti-individualistic, and antidemocratic (E. Gentile, 2003, 2004; G. Gentile, 2002; Gregor, 2001; Ioanid, 2005; Laqueur, 1996; Lederer, 1937; Schuman, 1934; Sternhell, Sznajder, & Asheri, 1994; Szaz, 1963; Wellhofer, 2003).

Although it may appear that the above litany of negations encompasses everything, fascism demanded cultural and ideological unity among all within the nation by forcing the creation of a new society, a new way of thinking, and a new man. Thus it was a totalitarian ideology. It was fiercely nationalist and jingoistic once in power, employing myth in order to stimulate nationalist fervor among its followers and seeking to eliminate all political opposition through violence.

Fascism came in different forms. The two most prominent were in fascist Italy and Nazi Germany. Fascist parties arose in other places, as well, but did not achieve the

same success as the National Fascist Party in Italy and the National Socialist German Worker’s Party or Nazi (short for National Socialist) Party did.

Studies of Fascism

Fascism has been widely researched by both political scientists and historians. From 1945 until the mid-1960s, the traditional approach to studying fascism tended to focus on its negative aspects and treat fascism as a trivial, reactionary ideology without its own historic uniqueness or substance (E. Gentile, 2005). Beginning in the mid-1960s and until the 1970s, empirical studies of fascism began generating new scholarly work that stressed fascism’s revolutionary, as opposed to its reactionary, characteristics (Griffin, 2005).

Emilio Gentile (2005) calls this phase of studying fascism the first period of renewal among three periods of renewal in the research. Empirical work began informing theory that replaced the traditional cursory interpretations of fascism. One of the first and most important scholars contributing to a more objective study of fascism was George L. Mosse, a Jewish intellectual who had suffered Nazi persecution himself (E. Gentile, 2005). There is not much disagreement over what fascism is not, but this new, more objective approach to studying fascism demanded that scholars also define what it is, its appeal, and how it

defines man and society (Linz, 1976). As a result, many scholars have attempted a definition of fascism, without consensus, but their attention to detail and precision is indicative of the change in the scholarly work.

In the 1980s, studies of fascism focused less on theory and more on the history of singular fascist movements, their politics, organization, and institutions. Differences among these regimes led some scholars to question whether a general theory of fascism was even appropriate. In the 1990s, a partial consensus began to emerge about the basic nature of fascism, along with a greater focus on fascist culture and ideology (E. Gentile, 2005). Sometimes expanding on previous literature, these studies covered class, civil society, and even rational choice perspectives (Wellhofer, 2003). In addition, the literature has become more compatible as the various studies began working from similar conceptual frameworks (Griffin, 2005), allowing for greater cooperation among scholars studying the fascist phenomenon.

Political Spectrum

Fascists were hostile to parties on the left, center, and right but most commonly, although not always, allied with those on the radical authoritarian right (Payne, 1995). Although it is common to place fascism on the right, there are differences among scholars, and even among fascists themselves, as to where fascism actually lies on the political spectrum.

On the economic political spectrum, fascism was not on the right. Capitalist-style competition was seen as destructive to the unity of the nation. Although some free market policies were not rejected outright, the ability of the state to interfere in economics with impunity and its increasing need for the war effort lead some to equate fascist government with Communist Soviet government (Lederer, 1937).

However, fascism also rejected socialism. Benito Mussolini wrote, "The socialists ask, 'what is our program?' Our program is to smash the skulls of the socialists" (Laqueur, 1996, p. 50). Whereas fascists claim that socialism accentuates class warfare and therefore a type of economic civil war within the state, fascism's aim is to reinforce class solidarity to strengthen the state (Laqueur, 1996). The difficulty with classifying fascism on the political axis arises because the ideology does not fit neatly into any category. Instead fascism sought to create a new culture and ideology independent from others (Payne, 1995), with the goal of replacing them all; hence its fierce opposition to all other ideological competitors.

Ideological Precursors of Fascist Theory

In the 20th century, several developments contributed to power that ideologies could exert. Bracher (1984) explains that never before in history did the legitimacy of a political

system feel such a need to justify its existence intellectually. Never had this justification had such sophisticated communications equipment at its disposal and never had regimes been so capable of manipulating public opinion.

Several strands of consciously irrational and illiberal political thought and historical trends contributed to the formulation of fascist thought. This chapter explores those ideological as well as cultural evolutions, their reasoning, and their prejudices. In so doing, much of the most important aspects of fascism are covered. Names of some of the people whose ideas led to fascist thought are mentioned although the list is by no means exhaustive. In addition, it should be noted that some of the people mentioned here, had they lived long enough, may not have actually approved of fascism and all that is associated with it.

Nationalism

Nationalist movements rejected rationalist thought, perceiving that it blunted nationalist sentiment with its atomization of society into individuals and cosmopolitan ideas. It was emotion and instinct that constituted reality, truth, and beauty, not rational thought. The nationalist believed that rationalism would ultimately destroy national activity (Sternhell et al., 1994).

The French politician Maurice Barres was one of the first to use the term *national socialism*. He believed that only emotion had real value and that real thought took place on the level of the unconscious. Therefore, attacking the unconscious with rationality divested the national organism of its substance. As a consequence, the welfare of the nation depended on the energy of the people. Rationalism and concepts of individualism were like a virus and would therefore contaminate the concept of the nation. The nation requires unity, and therefore a Marxist, liberal, proletarian, or bourgeois movement is antithetical to the idea of a nation. Enrico Corradini would later apply Barres's ideas and in 1910 would use the term *national socialism* to define Italian nationalism (Sternhell et al., 1994).

The reaction to natural rights theories and intellectualism took place in Germany in the forms of nationalism and romanticism. One of the best known German romantic nationalists was Johann Gottlieb Fichte. At the beginning of the 19th century, he argued for an independent German state and spoke of the cultural superiority of the German people. In contrast to social contract theory, the romantic nationalists saw the state as a living organism that survived through its national idea. The belief of Fichte and others was that a national consciousness had emerged from the concept of empire and the longing for an empire and that patriotism, no longer toward kingdoms but toward a German nation, became a historical force (Szaz, 1963).

National romanticism and the desire for the unification of German-speaking peoples in Europe was later expanded

on by Ernst Moritz Arndt to incorporate Prussian concepts of duty and a national will to power in order to create a unified people. This thought spread across Germany, with philosophers, writers, and educators introducing the concept of the German *volk*, or people, power with increasing popular support in Germany from the late 19th century until World War II (Szaz, 1963).

Pan-Germanism, in fact, influenced the war aims of German parties on both the right and the left during World War I. The nationalist ideas and racial beliefs of the superiority of the German people were intensely indoctrinated into the German army as well (Holborn, 1964). These concepts and ideas would later facilitate and influence Nazi empowerment, as well as Nazism.

Georges Sorel and Revolutionary Syndicalism

At the beginning of the 20th century, Marxists in France and Italy began questioning whether the theories of Marx were an actual reflection of reality. Proletarian revolutions were not occurring in countries where industrialization was most effective, in direct contradiction to what Marx had predicted. This realization of the failure of classic Marxism led to a split into two directions among European Marxists. In 1905, the Socialist Party formed and encompassed most socialists. The Socialist Party accepted the liberal democratic norms prevalent in Western European countries, with the objective of eventually changing the economic system through the democratic process (Sternhell et al., 1994).

Georges Sorel was a Marxist theoretician who represented another, more radical school of thought in the bifurcation of the European Marxists. Sorel did not reject capitalism and saw no difference between capitalist and Marxist economics. He believed that capitalism produced tensions between the classes that would lead to an all-out violent class struggle, which he advocated and which had been prognosticated by Marx. The problem according to Sorel was democracy. When conflict occurred between the proletariat and bourgeois classes, the democratic process allowed for compromise to diffuse the situation. If the democratic system could be destroyed, then the proletariat could be broken away from its alliance with the democratic socialists. Realizing this, the Sorelians *advocated the theory of revolutionary syndicalism* (Sternhell et al., 1994).

Besides the beliefs in the role of the market as an origin of tensions and overthrow of the democratic system, Georges Sorel introduced the doctrine of *social myth* into the syndicalist movement. Sorel believed that one of the advantages of introducing myth into politics was that myths were not subject to scientific criticism, and therefore doubt could not easily be introduced into the minds of the followers. In this sense, Sorel referred to Christianity and its use of the Second Coming as an effective myth. So he proposed the *myth of the general strike*, believing it an

update of Marx's revolution of the proletariat (Cohen, 1962). In this myth, the syndicates, or trade unions, would be the standard-bearers of Marx's revolution.

However, when the proletariat refused its role as the standard-bearer of the revolution, the Sorelians passed this task on to the entire nation. The result was a fusion, in both France and Italy, of the revolutionary syndicalists and the nationalists. The addition of nationalism contributed the *cult of a strong authority* to the syndicalist ideas. The revolutionary syndicalists were among the founders of Italian fascism and included Benito Mussolini (Sternhell et al., 1994).

Mysticism

Totalitarianism is the result of a revolutionary political movement's securing itself as the sole power in the nation and then proceeding to conquer society, seeking to politicize all existence according to its ideology. To accomplish this, the totalitarian regime would need to portray itself as a type of political religion through deifying the secular entity. This effort is facilitated through the use of myths (E. Gentile, 2005). The regime's justification to exist, therefore, could not be challenged because the myth provided an indisputable and indefinite source of legitimacy.

Myths were often extracted from history. Mussolini spoke of the return to the glory of the Roman Empire, and Hitler introduced the idea of the Third Reich as a new thousand-year empire (Koehl, 1960). In Italy, before the Fascists came to power, myth was already being employed. The Black Shirt militias were formed into units using terminology and organization based on those of the ancient Roman Empire (Payne, 1995).

Nazism's chief intellectuals, Alfred Rosenberg and Heinrich Himmler (head of the Schutzstaffel, or SS, which was the Nazi Party's personal and politically influential guard), as well as many national socialist historians, began using the history of the *Teutonic Order*, a small band of Germanic knights who had fought in the Crusades and existed for centuries in Germany. The importance of the Order in Nazi propaganda was its elitism. Only those dedicated and subordinate to a higher purpose, an idea not revealed to them, could be part of this Order. Rosenberg summoned the National Socialist German Worker's Party to be the German Order serving an "unknown god." Likewise, Himmler instilled into the SS a type of piety and worship of nature. On their belt buckles he had written, "My Honor Is Loyalty" (quoted in Koehl, 1960, p. 924).

In the 1850s, intellectuals in Germany began turning to the mystic racist idea of the superiority of the Nordic race in order to give themselves a belief in the future of the German nation, which at the time still had not formed, to the disenchantment of the nationalists (Szaz, 1963). Nordic superiority was the most heavily used myth in Nazi thought, replacing rationalism with racist mysticism. Racism and anti-intellectualism were the dominant tone, as

exemplified by one of the Nazis' favorite slogans, "We think with our blood" (quoted in Schuman, 1934, p. 211).

The use of irrational thought and mysticism by fascist movements does not deny that fascism has its own rationality. Fascism linked irrationality with rationality in the same way that religion links its supernatural ideas with its organizational institutions (Griffin, 2005). So the use of myth was a way of turning fascism into a political religion that its adherents would embrace in every way, politically, socially, economically, and spiritually.

Filippo Tommaso Marinetti and Futurism

Founded by Filippo Tommaso Marinetti, *futurism* was an art movement as well as a political movement. Futurism's political ideology was a form of radical nationalism advocated by young militant intellectuals in Italy. They embraced modern technology, youth, and violence and believed in man's dominance over nature. Futurists despised everything old, whether political or artistic, and sought its destruction. War was glorified, and their aggressive nationalism sought a greater Italy through the process of a cultural revolution and the development of a "New State" and a "New Man." Their political ideology often contradicted itself, though. Although virulently imperialist and supporting a militarist nation-state, they also praised cosmopolitanism and individualism, as well as libertarianism (E. Gentile, 2003).

Although the lack of coherence may have made their ideology harder to incorporate into fascism, Sternhell et al. (1994) posits that the common denominator among them, the revolutionary syndicalists, and the nationalists was their desire to destroy the dominant culture and replace it.

Giovanni Gentile and Actual Idealism

In 1921, Mussolini had written that fascism needed "a body of doctrine" if it were not to self-destruct (Gregor 2001, p. 33). According to Gregor (2001), the Italian intellectual Giovanni Gentile's *actual idealism* or *actualism* was such a doctrine, which G. Gentile, with the approval of Mussolini, infused into fascism. Actualism reiterated some of the themes expressed by the different ideological strands that led to fascism. For one, war was considered essential for the purpose of uniting the nation through shedding blood together. G. Gentile (1932/2002) wrote that during World War I, it was essential that Italy enter the war, and that it did not matter whether Italy had entered on the side of Germany or against it.

G. Gentile (1932/2002) defines fascism as a totalitarian ideology that does not concern itself only with politics but also with the thought and will of the nation. Individuals were social creatures, not isolated individuals having inalienable rights. Liberty exists only through the authority of the state and is manifested only as the liberty of the state rather than the individual. Therefore the state was not the result of a social contract but was instead a

fundamental part of human life. It was the social essence of individuals who shared a nationally defined consciousness (Gregor, 2001).

For the state to best represent the collective consensus of all citizens, the economy must reject the unbridled competition of capitalism, as well as the class conflict of socialism. Instead, G. Gentile (1932/2002) advocates a corporative nation wherein the state associates workers of a particular category with others in the same category into a type of union. However, these unions are not to work against one another but are rather to work harmoniously together as one national economic organism.

G. Gentile (1932/2002), an educated person himself as a professor of philosophy, reinforced anti-intellectualism in fascism. He wrote that the first among those who needed to be defeated included authors, cultivators of literature, and other academics, all of whom he called intellectuals. He opposed these intellectuals, not because he denied science, but rather because scientists believed the world existed independent of the mind. One of the tenets of actualism was that reality was only what was perceived by the mind and was therefore dependent on it. Gentile therefore rejects the positivist interpretation of the world and called it a "disease of reason" (quoted in Gregor, 2001, p. 28).

Biological Determinism

Although Italian Fascism and Nazism were similar in many respects, the German fascists succeeded in carrying out these principles more thoroughly. Mussolini, in order to justify Italian imperialism in Africa, championed the state as creating the nation. Hitler, in order to justify its claim to German-speaking areas in France and Czechoslovakia, claimed that the German nation was superior and that the state was its instrument (Cohen, 1962). But the main point of divergence between them was in the Nazis' *biological determinism*.

Linguistic and anthropological studies of the 19th century had revealed similarities in the languages of the people of Europe and central Asia. The assumption was that these languages had originated from a yet unknown common ancestor referred to by scholars as the *Aryans*. Many theories were developed to explain this finding, but that of the French count Arthur de Gobineau was to be the most significant for the development of Nazi biological determinism. Gobineau argued that Aryans had once been superior to all other races. However, they intermarried with various other races, diluting their purity and causing them to lose their superiority. Aryan blood was superior enough among the nations of northwest Europe, with Germany being the purest of all. While most of Europe gave no credence to this idea, it was embraced by Germany (Baradat, 1991).

Among those who embraced the Aryan myth was Gobineau's friend Richard Wagner, a very influential German composer in the mid-19th century. Wagner contributed to nationalism in Germany by emphasizing the

concept of the dynamic character of a nation, or its life-force of dynamism, and insisted on racism in politics and the teaching of anti-Semitism. In addition, Wagner introduced the idea of the king as a type of superhuman. From this perspective, the king spoke for the people, and therefore any constitutional limits on the king's power were interpreted as a humiliation or lack of confidence. In his last essay, perhaps disenchanted with the leaders of his day, Wagner had his idyllic superhuman ruler become a future leader who would unite the people, rejuvenate the national culture, and restore purity to the Aryan race (Szaz, 1963).

Houston Stewart Chamberlain, an intellectual and the nephew of British prime minister Neville Chamberlain, was a staunch advocate of Wagner's ideas and married Wagner's daughter Eva. He continued where Wagner left off. He claimed that the Aryan race had actually created all the other races but that their advances were negated through interbreeding. The only two exceptions were the German, who was Aryan and represented good, and the Jew, representing evil. Chamberlain concluded that if the Germans could remain racially pure, they would demonstrate their superiority by eventually conquering the world. This idea was extolled by Kaiser Wilhelm II, who then befriended Chamberlain, and by Hitler, who would incorporate Chamberlain's theories as the basis of Nazism (Baradat, 1991).

The philosopher Friedrich Nietzsche also had a significant influence on Hitler's thought. Nietzsche's theme that "might makes right" had an impact on Hitler. Nietzsche, in his writings, undertook a full assault on Christian and democratic values, which in his view protected the weak, thus preventing the eventual production of supermen. Szaz (1963) says that this denial of Christian morals by Nietzsche would relax the moral restrictions against the atrocities that occurred during World War II.

Hitler divided the world's people into three categories. The culture-creating race was the Aryans and included the English, Dutch, Scandinavians, and Germans, with the Germans being the most pure. All cultural achievements were the products of the Aryan peoples, and Hitler said that if the Nordic Germans were taken away, then all that would be left would be the "dance of apes." The second category was the culture-bearing races, which included the Asians, Latinos, and Slavs. According to Hitler, they could not create culture but could preserve it if they remained uncorrupted by inferior races. The culture-destroying races included Gypsies, Negroes, and Jews. They were responsible for the decline of civilizations and so, according to Hitler, deserved to die (Baradat, 1991).

Anti-Semitic legislation was eventually passed in Italy in order to convince Germany of its dedication to their alliance. According to Gregor (2001), Mussolini undertook an effort to make some form of racism and anti-Semitism a part of Italian Fascism in the summer 1938. However, Italian Fascism, unlike Nazism, did not contain any inherent racism. Fascism upheld the ideal of the nation-state even if

its history and culture were multiracial. Italy's Jewish community had been there since Roman times. So to be racist against Jews would attack the history of Italy. There were in fact a number of Italian Fascist Jews, and their numbers were greater than their ratio in the population. But an alliance with Hitler's Germany was necessary if Italy was to achieve its foreign policy goals. However, racist policies were not adopted wholeheartedly in Italy, as Fascist officials would intervene on behalf of the protection of Jews in many cases and the Italian military command even provided protection to Jews as well.

Historical Context

Fascist movements were anticapitalist movements in that they sought to reshape the capitalist economic order into one that still sought economic growth while eliminating the tensions between employer and employee caused by industrialization and exacerbated by dire economic prospects after World War I (Fletcher, 1979). Private property was still allowed, but the democratic framework within which it operated was eliminated and placed under state control to preclude private enterprise from contradicting the wishes of the state (Cole, 1941). Fascism's goal in doing so was to rectify what it saw as capitalism's fragmentation of society into self-interested individuals and antagonistic groups that it believed dehumanized people's relationships (Sternhell et al., 1994).

History showed that fascist movements were successful only in democratic societies, where they were allowed to roam freely and spread their ideas. When they tried to form inside an authoritarian regime, they were always crushed. Where they were successful, their success was due to a lack of support for the democratic regime (Laqueur, 1996). In most democratic regimes, fascists formed only fringe parties and never gained real influence. The two countries where they did find success and power were Italy and Germany.

Italy

Italy experienced rapid economic growth at the beginning of the 20th century, along with increasing nationalism and a desire for empire. The government, headed by Givoanni Giolitti, responded to this sentiment by invading and capturing Libya from the Turkish Empire. However, this conquest failed to quench the nationalists' imperial thirst (Payne, 1995). This thirst for imperial possessions determined which side Italy would fight on in World War I. It fought with the side that it felt would grant it the greatest amount of territory. But after the war was over and the Allies had won, Italians felt that they had not been compensated adequately for the alliance. The result was that the nationalists began to denounce the political leaders who accepted the Allies' terms and labeled the victorious outcome of the war the "truncated victory" (Payne, 1995).

The trenches of World War I tied down millions of soldiers in stationary combat for long periods of time. A camaraderie and collective consciousness developed and was made more potent by the shared suffering (Payne, 1995). After the war, these veterans would return home to find few prospects. The process of modernization had been accompanied with instability, unemployment, and inflation, which plagued many European countries. The Italian kingdom was in a state of malaise, with massive social and economic problems (Baradat, 1991), as well as opportunity for those willing to exploit it. Although developed in France, it was in Italy where the revolutionary syndicalists became a significant political force. In 1914, the revolutionary syndicalists, nationalists, and futurists found the perfect setting to allow them to transform their ideologies into a historical force (Sternhell et al., 1994).

Germany

Germany was in bad shape following World War I, both economically and socially. People rejected the peace that had long existed since the end of the Napoleonic Wars in 1815, the liberal cultural synthesis, and Germany's contemporary leadership. Political life became brutalized, and the government increased its control over society while curtailing civil liberties. Fueling this was rampant hyperinflation and chaotic social conditions, at the same time that the inception of mass media allowed those with extreme solutions an outlet for their ideas (Payne, 1995). The Nazi party blamed all Germany's woes on the Jews. The reasons for Germany's loss in World War I were also the Jews' fault because, according to the Nazis, they were responsible for the establishment of parliamentary democracy and what Nazis called the "Jew republic" of the "November criminals" of 1918. The signing of the Treaty of Versailles was seen as a stab in the back of the German military. Likewise, all of Germany's problems following the armistice were attributed to the Jews (Schuman, 1934).

The Treaty of Versailles had unjustly assigned Germany all the blame for the war and sought excessive punitive actions in the form of territorial concessions and inordinate reparations. Like Italy, Germany experienced very high unemployment and inflation. The difficulty of the Weimar Republic in solving these problems led many to turn to extreme political movements such as the Nazis. An incipient economic recovery was rudely interrupted by the Great Depression, which struck in 1929, bringing about the conditions for the collapse of the democratic regime (Baradat, 1991).

Leaders

All fascist movements required a charismatic leader and developed a cultlike following of that leader. Mario Palmieri lays out "The Hero as Leader" in his work *The*

Philosophy of Fascism (cited in Cohen, 1962). He defines the hero as he who can rediscover the greatest of truths. The true hero will be sincere and courageous and will believe in his own destiny. Palmieri adds that in addition to the virtuous human traits above, the hero will possess a mystic power of intuition that enables him to obtain immediate knowledge of the truth. According to Palmieri, Mussolini was the hero and expressed what was in everyone's hearts, but only in the role of supreme leader would he be able to change the world (cited in Cohen, 1962).

Palmieri's Nazi counterpart in defining the role of the leader can be found in theoretician Ernst R. Huber. Huber explained that the true will of the people could not be given by democratic means but could be conveyed only through the Fuhrer. The Fuhrer's will is not his individual will, according to Huber, but rather the collective will of the nation, which is embodied within the Fuhrer. The state therefore has no inherent authority but rather derives its authority from the Fuhrer to apply the national will. The Fuhrer has no political constraints, but Huber asserts that he is not self-seeking and will exist to apply the true will of the people (Cohen, 1962).

Benito Mussolini

Benito Mussolini was born in 1883 to a mother who was a schoolteacher and a blacksmith father who was also a Socialist. He was named after Benito Juarez, the former Mexican president. He violently assaulted fellow students on several occasions but later became an elementary school teacher. While staying in Switzerland, he became a Socialist and would later write for different Socialist papers, eventually becoming the editor of *Avanti* (Forward), the Socialists' official newspaper. He became one of the top leaders of the Socialist Party at the age of 29 (Payne, 1995).

Although he was a Socialist, Mussolini rejected egalitarianism and was heavily influenced by the theories of Georges Sorel. Most significantly, Sorel's writings inculcated in Mussolini the idea that great historical events were set in motion by the initiative and leadership of small groups of people. Seeing the widespread war fever that overtook many European nations at the beginning of World War I, Mussolini noticed the appeal of nationalism and opportunistically went against the Socialist stance of neutrality by asserting that Italy should enter the war. This action cost him the editorship of *Avanti* and his membership in the Socialist party. Mussolini would again act on opportunism when a secret pact with the Allies to grant Italy territorial concessions for joining the war effort was not honored. The resulting confusion and postwar social, political, and economic turmoil gave Mussolini the opportunity to found the Fascist Party (Cohen, 1962).

Unable to meet with much success at the polls, Mussolini began brazenly promoting the idea that he would take the government over by force. On October 27, 1922, although outnumbered and inferior in strength to the

police, an army of thousands of profascists began marching on Rome. The Fascists took over many of the police stations without having to use much violence even though Mussolini's Black Shirts, who were thugs who used intimidation and violence in the name of the Fascist Party, were ready to use terrorism. The chances they would succeed in a coup against the government were unlikely, and the army was ready to fight the Black Shirts if the King requested.

Armed with clubs, the Fascists arrived outside of the city on October 28. However, King Victor Emmanuel refused to act against Mussolini and instead invited him to lead a new parliamentary coalition. King Emmanuel could have easily stopped Mussolini, but for some reason did not. Perhaps he did not have the courage to act. Or maybe he believed that fascism was a good direction for the country to go. Whatever the reasoning, this decision placed Mussolini in control of Italy's destiny and made fascism a historically significant ideology in the world (Payne, 1995).

Mussolini (1968) wrote down his ideas on fascism some time after taking power. In 1935, he wrote that liberalism had arisen as a reaction to absolutism but had outlived its function once the state became the expression and will of the people. Liberalism, according to Mussolini, tried erroneously to elevate the importance of the individual over the state, but it was the state that expressed the true conscience of the individual. Mussolini proclaimed that it was his job to reassert the right of the state.

Mussolini (1968) explains that by following the spiritual attitude of fascism, one will see the common bond of tradition and mission of the nation and of the individual, which will suppress one's instinct to live, thus allowing one to break free from the constraints of time and space through self-sacrifice. By renouncing self-interest through death itself, one can accomplish a spiritual existence. So if the followers of fascism are willing to sacrifice themselves for the cause, they can achieve a kind of immortality, giving the ideology a religious zeal.

Adolf Hitler

Adolf Hitler was born to a minor customs official and developed a strong sense of German nationalism at a young age. At first trying to become an artist, Hitler went to Vienna in 1906 but experienced only rejection by the city's leading art schools (Baradat, 1991). While in Vienna, Hitler encountered the ideas of pan-Germanism from Georg von Schonerer and the Christian Social movement under Vienna's Mayor Karl Lueger, which gave him his first encounter with anti-Semitism, as well as its popularity. Hitler delved deeper into theoretical anti-Semitism by reading pamphlets that were created by a former monk called Lanz von Liebenfels. Hitler most likely read Wagner's racist writings as well (Holborn, 1964).

After suffering a poison gas attack as a soldier in World War I and sitting out the rest of the war, Hitler joined other Germans claiming that Germany had not lost the war but

rather had been betrayed by a Jewish conspiracy. He later joined the Nazi party, which had only six other members at the time, but this small group became the core of the National Socialist German Worker's Party. He quickly rose to the leadership position of the party as everyone recognized him as having leadership qualities. After attracting new members, including some important military people, Hitler was inspired by Mussolini's March on Rome to attempt his own coup, which failed. The Beer Hall Putsch, as it was called, resulted in his imprisonment, but because of powerful sympathetic allies, he ended up serving only a year in prison (Baradat, 1991).

While imprisoned, Hitler (1939) wrote his ideology down in *Mein Kampf* (My Struggle). Mostly it is an extremely long rant. However, the ideas broached in it form the basis of Nazi thought. In *Mein Kampf*, Hitler chides Jews with numerous invectives. He details his encounter with Jews in Vienna and how he went from seeing them as equals to uncovering their conspiracies in the form of both Zionism and liberalism and eventually having nothing but vitriol for them. Hitler explained that the world will be ruled either by liberal democracy, wherein the numerically superior races would reign, or by the law of natural distribution, whereby the most brutal nation would reign supreme through war, which Hitler would set out to initiate once in power.

One of the fundamentals of his thinking was the idea that life was an eternal struggle in order to dominate others. Hitler said that only force rules and that humanitarianism was nonsense. Struggle was the prerequisite needed for human development and progress (Cohen, 1962). Hitler declared that it is not people who move history but rather races that do. Using Gobineau's theories on race, Hitler proclaims that the German people must purify themselves by eliminating inferior races. The Nordic, or Aryan, race is the most superior, and therefore the greatest application of resources should be used to enhance the breeding of this race. If agriculture cannot sustain that effort, then Hitler argues that this is a justification for expelling inferior races from German lands through war or even annihilation (Holborn, 1964).

Hitler believed that his own ideology possessed the principles for rebuilding Germany and establishing its supremacy, as opposed to liberal democracy's principles, which he viewed as weak. But before this struggle could be undertaken abroad, it first had to be won in Germany. This required the defeat of liberalism, socialism, and communism, as well as the implementation of a totalitarian ideology immune to foreign propaganda (Holborn, 1964).

The Great Depression gave the Nazis increasing influence as voters became disillusioned by the failures and indecisiveness of their political leaders. Believing they could control Hitler, conservative politicians persuaded president Paul von Hindenburg to appoint Hitler as chancellor. However, Hitler outmaneuvered them by having the Nazis burn down the parliament building, blaming the

communists, and arresting his main opponents before the elections, resulting in Nazi dominance of the Reichstag. When von Hindenburg died the following year, Hitler took power and outlawed all opposition (Baradat, 1991).

Hitler had a profound contempt of generals, the bourgeoisie, and professors. None of these groups ever merited the same hatred as his hatred of Jews, though. Although anti-Semitism was popular in Germany, his was genuine and not just the result of opportunism. During the anti-Jewish pogroms of April 1, 1933, and November 8–9, 1938, known as *Kristallnacht* (Crystal Night), or the Night of the Broken Glass, the violence was carried out in public. However, even when violence against Jews would not be popular, as would be the case with extermination camps, Hitler was not swayed. He ordered the extermination of European Jewry in 1942 anyway, but in secret in order to avoid the likely disapproval of the German people (Holborn, 1964). At that fateful meeting in Wannsee, a suburb of Berlin, on January 20, 1942, 14 people, half of whom had PhDs, met, and the “final solution” to the Jewish problem was proposed: After using them for their labor skills, they should be eliminated (Chodoff, 1997). So commenced the Holocaust, Hitler’s infamous legacy.

Other Fascist Movements

A problem with generic definitions is that the essence of fascist movements is national, not international, and therefore national differences among the fascist movements of different nations are unavoidable. So questions arise as to whether these differences disqualify a movement from being considered fascist. This chapter avoids the debates over whether these movements are actually fascist and presents instead a cursory introduction to other movements that are most commonly thought of as being fascist.

Imperial Japan during World War II did not have a single mass party, no dictator seized power, and no totalitarian ideology became dominant, and therefore it was very different than the Fascists of Europe. The New Order Movement, which failed to gain power, was a fascist movement in Japan, however. Its reforms were modeled on the institutions of Fascist Italy and Nazi Germany and sought to build an economically strong Japan while at the same time eliminating the tensions of industrialization, as fascists had tried to do through the corporative economic system (Fletcher, 1979).

During World War II, many fascist movements sprang up across Europe. Among these were the Arrow Cross in Austria-Hungary, the Falange in Spain, and the Rexists in Belgium (Laqueur, 1996). A tactic that all these movements had in common was violence. They all used terrorism from below when out of power and terrorism from above once in power. Almost always, this violence was carried out collectively rather than by individuals. When in opposition, they would organize gangs in order to break

up their adversaries’ assemblies by beating them up or sometimes killing them (Laqueur, 1996).

One fascist movement that heavily relied on violence and mysticism was the Iron Guard in Romania, or Legion of the Archangel Michael, known also as the Legionary Movement. It employed a brutally intense cult of death against its adversaries. Its founder, Corneliu Zelea Codreanu, believed that the nation constituted all Romanians, alive, dead, and those yet to be born. Followers believed that their mission was God’s secret will, which was organically fused with the nation, the state, the king, and the Legionary Movement. This destiny had to be protected from any outside influence, physical or cultural (Ioanid, 2005).

The skinheads, despite their apparent admiration for Adolf Hitler, the swastika, and the Nazis, are not considered a fascist movement. Their dress, music, and tastes are a mix of different cultures, in contradiction to Fascist nationalist doctrine. Laqueur (1996) says that they do not have the knowledge, motivation, or discipline to be considered of any use to neofascist elements.

Contemporary established fascist parties appear to have little in common with those of Nazi Germany and Fascist Italy in that they observe the democratic rules and norms. However, Laqueur (1996) warns that this is because they are all relatively weak and so must be cautious in the face of strong regimes, but that such may not be the case if they were to gain power.

Conclusion

Fascism and National Socialism developed from several ideological trends: nationalism, Revolutionary Syndicalism, Futurism, Actualism, and in the case of Nazi Germany, biological determinism. What these ideologies had in common was that they despised the atomization of society into individuals, brought about by liberal democratization and capitalist-led modernization, and they extolled the concept of the organic nation-state in order to establish more cohesive collective ties among people. These ideologies maintained that an integration of all society would produce these results, but all nonconformist elements had to be co-opted or exterminated. In the case of the racist policies of Nazi Germany, the former was not an option for inferior races.

Fascists believed that liberal democracy was weak and would never accomplish these goals, so they advocated a totalitarian government headed by a strong leader. World War I had desensitized much of Europe to extreme violence, and the Great Depression had left the masses economically insecure. Moreover, the young liberal democracies in Germany and Italy proved ineffective at solving these problems. Economically, politically, and spiritually, Italians and Germans were seeking salvation, and so the Fascist and Nazi parties filled the vacuum. Mussolini and Hitler also

brought with them their violent plans for war, with grave consequences for the rest of the world.

References and Further Readings

- Baradat, L. P. (1991). *Political ideologies: Their origins and impact* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Bracher, K. D. (1984). *The age of ideologies: A history of political thought in the twentieth century* (E. Osers, Trans.). New York: St. Martin's Press.
- Carsten, F. L. (1982). *The rise of fascism*. Berkeley: University of California Press.
- Chodoff, P. (1997). The holocaust and its effect on survivors: An overview. *Political Psychology*, 18(1), 147–157.
- Cohen, C. (Ed.). (1962). *Communism, fascism, and democracy: The theoretical foundations*. New York: Random House.
- Cole, T. (1941). National Socialism and the German Labor Courts. *Journal of Politics*, 3 (2), 169–197.
- Fletcher, M. (1979). Intellectuals and fascism in early Showa Japan. *Journal of Asian Studies*, 39(1), 39–63.
- Gentile, E. (2003). *The struggle for modernity: Nationalism, futurism, and fascism*. Westport, CT: Praeger.
- Gentile, E. (2005). Fascism, totalitarianism and political religion: Definitions and critical reflections on criticism of an interpretation (N. Belozentseva, Trans.). In R. Griffin (Ed.), *Fascism, totalitarianism and political religion* (pp. 32–81). New York: Routledge.
- Gentile, G. (2002). *Origins and doctrine of fascism* (J. Gregor, Trans.). New Brunswick, NJ: Transaction. (Original work published 1932)
- Gregor, J. (2001). *Giovanni Gentile: Philosopher of fascism*. New Brunswick, NJ: Transaction.
- Griffin, R. (Ed.). (1998). *International fascism: Theories, causes and the new consensus*. London: Arnold.
- Griffin, R. (Ed.). (2005). *Fascism, totalitarianism and political religion*. New York: Routledge.
- Hitler, A. (1939). *Mein kampf* (J. Murphy, Trans.). London: Hurst & Blacket.
- Holborn, H. (1964). Origins and political character of Nazi ideology. *Political Science Quarterly*, 79(4), 542–554.
- Ioanid, R. (2005). The sacralised politics of the Romanian Iron Guard. In R. Griffin (Ed.), *Fascism, totalitarianism and political religion* (pp. 125–159). New York: Routledge.
- Koehl, R. (1960). Feudal aspects of National Socialism. *American Political Science Review*, 54(4), 921–933.
- Laqueur, W. (1996). *Fascism: Past, present, future*. New York: Oxford University Press.
- Lederer, E. (1937). The economic doctrine of National Socialism. *Annals of the American Academy of Political and Social Science*, 191, 219–225.
- Linz, J. J. (1976). Some notes toward the comparative study of fascism in sociological historical perspective. In W. Laqueur (Ed.), *Fascism: A Reader's Guide* (pp. 3–121). Berkeley: University of California Press.
- Mosse, G. L. (1964). *The crisis of German ideology: Intellectual origins of the Third Reich*. New York: Grosset & Dunlap.
- Mussolini, B. (1968). *Fascism: Doctrine and institutions*. New York: Howard Fertig.
- Nolte, E. (1982). *Marxism, fascism, cold war*. Atlantic Highlands, NJ: Humanities Press.
- Payne, S. G. (1983). *Fascism: Comparison and definition*. Madison: University of Wisconsin Press.
- Payne, S. G. (1995). *A history of fascism, 1914–1945*. Madison: University of Wisconsin Press.
- Schuman, F. L. (1934). The political theory of German fascism. *American Political Science Review*, 28(2), 210–232.
- Sternhell, Z., Sznajder, M., & Asheri, M. (1994). *The birth of fascist ideology: From cultural rebellion to political revolution* (D. Maisel, Trans.). Princeton, NJ: Princeton University Press.
- Szaz, Z. M. (1963). The ideological precursors of National Socialism. *Western Political Quarterly*, 16(4), 924–945.
- Wellhofer, E. S. (2003). Democracy and fascism: Class, civil society, and rational choice in Italy. *American Political Science Review*, 97(1), 91–106.

76

MARXISM

MICHAEL D'AMORE

Sussex County Community College

JOHN T. ISHIYAMA

University of North Texas

Socialist political philosophy has existed since the beginning of recorded history. It has also taken on a great number of forms since its emergence in antiquity. However, no theory of socialism has had a greater impact on the modern world than the philosophy constructed by the 19th-century German thinker Karl Marx. Marx's theory of socialism originated from (and was a direct response to) the capitalist mode of production. Marx, particularly, focused on the relationship between capitalism as an economic system and industrial development in Western Europe during the middle of the 19th century. Along with his lifelong collaborator, Friedrich Engels (1820–1895), Marx wrote several epic volumes that impacted almost all Western political thought from his time through the present. Some of Marx's most influential works, such as the first volume of *Capital* and *The Communist Manifesto*, were published during his lifetime. However, many of his significant writings, such as two subsequent volumes of *Capital*, the *German Ideology* and the *Economic and Philosophic Manuscripts*, were published posthumously. In those volumes and in many more pieces, Marx developed an analysis of industrial capitalism that was both complex and comprehensive. After his death in 1883, Marx's theory was repeatedly expanded on by devotees and detractors alike. Whether they offered a new interpretation of a particular aspect of Marxian thought or a rigorous critique of his ideas, all those who responded to Marx ensured that his ideas will continue to live far beyond his corporeal existence. More than a century after his death,

Marx remains the unequivocal “father” of modern socialist thought.

To understand Marxism and its emergence, one must have some sense of the context in which it emerged, as well as of Karl Marx the man. It was a combination of his own experiences, the philosophical influences on his work, and the social and economic context of the 19th century that led to the emergence of one of the most powerful philosophical and ideological influences of modern times.

The Life and Times of Karl Marx

Karl Heinrich Marx was born in Trier, Germany, on May 5, 1818, to Hirschel and Henrietta Marx. His father was one of the most respected lawyers in the city, a man who had converted from Judaism to Protestantism in order to keep his job. The young Marx grew up in a comfortable middle-class household and led a fairly uneventful life. At 17, he enrolled at the University of Bonn to study law. At Bonn, he spent a great deal of time “socializing” and running up rather large debts from his adventures at local beer halls. He also became engaged to Jenny von Westphalen, the daughter of Baron von Westphalen, a prominent member of Trier society. When Marx's father found out that Karl had been wounded in a duel, the elder Marx insisted that his son withdraw and enroll at the more “sedate” University of Berlin, in the Prussian Empire (Wheen, 2002).

At the University of Berlin, professor Bruno Bauer (1809–1882) introduced young Marx to the writings of the German philosopher Georg Wilhelm Friedrich Hegel (1770–1831) and his philosophy of *dialectical idealism*. Bauer also introduced Marx to atheism and other radical political opinions that got Marx into trouble with the authorities. Marx was especially impressed by Hegel's theory that a thing or thought could not be separated from its opposite. For example, the slave could not exist without the master, and vice versa. Hegel argued that unity would eventually be achieved by the equalizing of all opposites, by means of the *dialectic* (logical progression) of *thesis*, *antithesis*, and *synthesis*. The ultimate thesis was the "truth." This was Hegel's theory of the evolving process of history and the ideals that motivated history. Marx joined the Young Hegelian movement, which fiercely criticized both the Prussian aristocracy and its opposition (McLellan, 1973).

Following the death of his father (and source of financial support) in 1838, Marx decided to earn a doctorate and become a university professor. However, after completing his doctoral thesis at the University of Jena, a thesis which was a fierce critique of spiritualism and which laid out the basis of *materialism*, the idea that material reality produces thought in humans and not the other way around, Marx was unable to find a teaching position (largely because of his radical anti-Prussian views). In 1842, he found a job in Cologne as the editor of a newspaper, *Rheinische Zeitung*, which opposed the Prussian attempt to dominate the West German principalities. As editor, Marx wrote a number of editorials that compelled the local government, under pressure from the Prussians, to close the paper. Marx quickly married his fiancé Jenny and then emigrated to France, arriving in Paris at the end of 1843 (Mehring, 2003; Wheen, 2002).

In Paris, Marx made contact with several noteworthy radicals, including the exiled Russian anarchist Mikhail Bakunin (1814–1876), the idealist anarchist Pierre J. Proudhon (1809–1865), and Marx's most important collaborator, Friedrich Engels, the son of a wealthy German industrialist. In Paris, Marx and Engels decided to work together, bringing to the table different skills: Marx was best at conceptualizing and abstraction, and Engels was better at communicating abstract concepts to a mass audience. Thus began a mutually beneficial lifelong partnership.

In 1844, the authorities expelled Marx. He moved his family to Brussels, Belgium, where he remained until 1847. Engels subsequently moved to England, where Engels's family had cotton spinning interests in Manchester. Marx had already published several works that outlined his theory of materialism and its impact on the development of history and predicted the collapse of capitalism. In Brussels, Marx joined the Communist League, a group of German émigrés with its center in London. Marx and Engels became the major theoreticians of the organization,

and at a conference of the League in London at the end of 1847, Marx and Engels were commissioned to write the program for the organization: *The Communist Manifesto*.

The Communist Manifesto was published immediately before the Year of Revolutions, 1848. These revolutions were a series of political upheavals throughout Europe in the spring of 1848. Essentially it was a revolutionary wave that began with the French revolution of 1848 and then spread rapidly throughout Europe. Although most of the revolts were put down very quickly, a significant amount of violence occurred, with tens of thousands of revolutionaries executed.

The causes of upheaval were many, but one of the major factors was the rise of industrial capitalism in Europe and the rapid urbanization of the population that accompanied industrial expansion. Early capitalism had led to rapid economic expansion but was also accompanied by the widespread misery of the working classes. Unemployment, poverty, and the lack of a political voice via the right to vote all contributed to the beginning of the upheaval. Although the revolution in France had started as a protest movement led by the middle classes against the Orleans monarchy of Charles X, the last Bourbon king, it quickly became an uprising of the working classes in the cities.

Early in 1848, Marx moved back to Paris when the revolution first broke out and then on to Germany, where he founded, again in Cologne, the *Neue Rheinische Zeitung*. The paper openly opposed the Prussian autocracy and pressed for revolt. The paper was suppressed, and Marx fled to London in 1849 to avoid arrest, an exile that was to last the rest of his life.

In the early period of his exile, Marx was quite optimistic about the prospect for another major revolutionary upheaval that would destroy capitalism and all its evils. He rejoined a resuscitated Communist League in London and wrote two pamphlets that argued that another revolution was imminent, the *Class Struggles in France* and the *18th Brumaire of Louis Napoleon*. As the years passed, he became more interested in the study of political economy in order to understand what led to the conditions for revolution. He spent the next years working in the British Museum and living in abject poverty in a three-room flat in the Soho section of London with his family. He had a total of six children and depended almost entirely on gifts from Engels, whose family business in Manchester was doing quite well. He also worked as foreign correspondent for the *New York Daily Tribune* at this time (Barnett, 2009).

Despite all his problems Marx continued to work, and in 1867, the first volume of his greatest work, simply entitled *Capital*, was published. The volume is a detailed analysis of capitalism and how it created the conditions of abject poverty and worker alienation. Marx also deals with the issue of revolution, arguing that capitalism creates the conditions for its own destruction. In 1871, Marx began working on the second volume of *Capital*. He had been

encouraged by the formation of the Paris Commune in March 1871 and the abdication of the French dictator Louis Napoleon but became despondent after the revolt collapsed. Volume 2 was never finished as Marx's health and his wife's deteriorated. Jenny Marx died in 1881, and Marx's eldest daughter died in January 1883. Karl Marx died 2 months later, on March 14, 1883.

Marxism

From Marx's thousands of pages of writing (much published only after his death), some fundamental themes emerge. First, it is important to note that Marx built his theories on several assumptions that were prevalent in economic thought of the time. The first was the *labor theory of value*. The labor theory of value is a major pillar of traditional Marxian economics, which is quite apparent in Marx's masterpiece, *Capital*. The basic claim is rather straightforward: The value of a commodity is determined by the amount of labor that is invested into that commodity. For example, a primitive axe made of vines, wood, and a stone is more valuable than its component parts because of the labor invested in it. Or, if a pair of shoes takes twice as long to make as a pair of pants, then shoes are twice as valuable as pants, regardless of the value of the physical inputs. Although this theory has been disproven, early economists such as Adam Smith (1723–1790) and David Ricardo (1772–1823), who influenced Marx's thought, were proponents of this idea. The theory meant that because all value was created by labor, capitalism stripped the producers of their humanity by extracting "surplus value" for the benefit of the capitalists.

Another prevalent assumption in economic theory of the time was the *iron law of wages*. Although the idea is most closely associated with Ferdinand Lassalle (1825–1864) and Thomas Malthus (1766–1864), David Ricardo was said to subscribe to it. According to Lassalle, wages in capitalism are reduced to the *cost of reproduction*, or the amount required so that labor can physically reproduce itself by having children. This is because competition between firms requires that capitalists reduce the costs of production to be competitive. Because value is created solely by labor (according to the labor theory of value), then the primary cost of production is the cost of labor, or wages. Over time there is pressure to reduce wages to the minimal subsistence level, or the cost of reproduction. This idea was to play an important part in Marx's argument regarding the inevitability of the collapse of capitalism.

Based on these common economic assumptions of the period, Marx fashioned a comprehensive view of the evolution of human history that included several key elements. The first is his view of human nature, the second is related to his views on economic determinism and historical and dialectical materialism and his explanation of capitalism's

inevitable collapse, and the third is the role of professional revolutionaries in facilitating the revolution.

Human Nature

Marx's theory of socialism originated from his unique perspective on human nature. Unlike most of the prominent Western political theorists before him, Marx did not adopt an essentialist conception of human nature. Rather, he and Engels asserted in *German Ideology* that at any given time in human history, the "natural" condition of humankind was significantly influenced by the material and social conditions that were dominant at that moment (Tucker, 1978). From this perspective, humans are not naturally born with anything at all (in direct contrast with many of the liberal theories of the 18th century that spoke of natural inalienable rights). In particular, these material or social conditions affect all that we do as human beings.

It is important to note that Marx did not think that humans were merely passive reflections of their environments (as would be the case if one were to argue in favor of nurture as opposed to nature). Indeed, as other contemporary thinkers associated with the positivist movement of the 19th century (especially Auguste Comte, 1798–1857), Marx believed that humans had the ability to shape and change their material conditions. As a result, Marx claimed that humans were participants in the crafting of their own consciousness rather than simply passive blank slates whose nature changed with arbitrary changes in material conditions (Tucker, 1978).

Although Marx would agree in general that human beings were not born with anything, he believed one impulse was natural to humans. Like all animals, humans confront their surroundings as they find them and then alter the material world through their productive capacity. Humans are unique, however, because they are the only animal conscious of their own productive acts and have a natural desire to produce what they can imagine. Therefore, Marx claimed, the symbiotic relationship between human consciousness and the given material conditions, at any point in history, becomes established by conscious human action (Tucker, 1978).

Economic Determinism and Historical and Dialectical Materialism

Based on his assumptions regarding human nature (that human beings are naturally economic beings), Marx argued that everything that human beings create therefore has some economic purpose. Thus, everything—religion, culture, laws, government—is designed for particular economic purposes, generally to keep the dominant class dominant. For instance, Marx noted that most laws made by the state were meant to protect property, an instrument

by which one class rules another. In most of their writings, Marx and Engels seem to see the state as a neutral tool, much like a weapon (Evans, 1975). Similarly, religion has an economic purpose. As Marx notes, it convinces the oppressed of a better life after the current world (as long as they are obedient), thus making the oppressed accept their condition. “Religion is the sigh of the oppressed creature, the heart of a heartless world, and the soul of soulless conditions. It is the opium of the people” (Marx, 1844, para. 3).

Another important element of Marx’s theory is that history moves in distinct stages, and what causes movement from one stage to another is conflict, particularly class conflict. Here Marx draws on the notion of Hegelian dialectics. For Marx, class was defined by an individual’s relationship to the means of production. *Means of production* referred to those things that are used to produce other things. Thus, for example, land, water, and buffalo were the means of production for the Plains tribes of North America. These resources were used to produce other goods (such as shelter and tools). In Rome, the primary means of production were land and slaves. In modern industrial society, the means of production are the machines and factories used to produce other products that are then consumed. Class is determined by the extent to which people own most, some, or little of the means of production, or by their relationship to the means of production. It is generally conflict over control or access to the means of production that drives history.

For Marx, history is driven by the never-ending, cyclical process of humans’ acting on their material conditions, altering those surroundings, and, in turn, being affected by a newly generated set of material conditions. Inspired by Hegel’s distinctive theory of history and idealist philosophy, Marx postulated that human social and political development are advanced through conflict between antithetical class forces. Marx made a major departure from Hegel, however, on the nature of this conflict. Marx is said to have “stood Hegel on his head” by claiming that it was conflict rooted in the material conditions of existence that drove history, and not conflict over antithetical ideas, which Hegel asserted was the principal mover of human history. Thus Marx distinguished his own form of dialectic as *dialectical materialism*, in contrast to Hegel’s dialectical idealism.

Marx examined the dominant material conditions at various moments of human history and stated that each set of dominant conditions bred a set of conflictive conditions. In the hands of human beings, these contradictory conditions contributed to conflict; at times, this conflict became so deep and irresolvable that it transformed human development in profound ways. Marx asserted that human beings drove this process by acting collectively and particularly as members of an economic social class. As a result, for Marx and Engels, history moved in distinct stages or *epochs*, and within each epoch, one could find the

contradictions (or class conflicts) that would pave the way to the next stage. Marx identified the following stages:

- Primitive communism
- Slave society
- Feudalism
- Capitalism
- Socialism and communism

Unlike earlier liberal democratic theory, which held that there had been a time in human history when humans did not live in a society (or the so-called state of nature), Marx argued that humans had always lived in some kind of society. The first of these societies he called *primitive communism*. Although Marx is associated with this term, primitive communism was most fully elaborated by Engels (1884), who thought of it as a period when the collective right to basic resources, egalitarianism in social relationships, and the absence of authoritarian rule and hierarchy all existed. This stage was characterized by a society much like the tribal communities of the North American Plains. Although humans possessed personal items (their clothes, some tools, etc.), there was no sense that individuals owned the major means of production—the land, the water, the buffalo, and so forth. Without private property (in this sense of ownership of the means of production), there were no classes to speak of. Since this was a classless society, it was communist. What made it primitive was the very low standard of living and the great dangers facing tribal members.

Eventually, primitive communism gave way to the next stage of history, *slave society*. Although Marx and Engels are not clear as to how primitive communism collapsed, there is a suggestion by Engels (1884) that it was a “natural” development. In other words, someone somewhere inevitably claimed a particular piece of land or a particular herd of cattle. This claim created the basis of the haves versus the have-nots, or class contradictions. Slave society was in many ways the first epoch with class contradictions. In slave society, the principal means of production were land and slave labor, as was the case in Rome. Wealth in slave societies was defined in terms of land ownership and slave ownership. In such societies, there were classes: those who owned most of the land and slaves (or most of the means of production), such as the large landholding patricians of Rome; those (such as artisans) who owned some of the means of production; and those who owned nothing, not even themselves (slaves). Societies such as Rome were rocked by internal conflicts among these classes for control over the means of production (such as the slave revolt led by the gladiator Spartacus in the 1st century CE). Eventually these conflicts led to the demise of slave society and the emergence of feudalism.

Feudalism, like slave society, was characterized primarily by agricultural production controlled by large estates of landholding nobles. However, unlike slave society, primary

labor was based, not on slavery, but on peasant and serf labor. Although serfs were legally bound to land and could not freely leave, unlike slaves, who were property, serfs owned themselves. In feudalism, there were also other classes, particularly the merchants, or the early bourgeoisie. The early bourgeoisie, unlike the landholding nobility, derived their livelihood from the control of trade (such as ships, transport) and finance. With the expansion of trade routes east and west, the European bourgeoisie grew in economic status and demanded political power as a result.

The Rise and Collapse of Capitalism

Ultimately, the bourgeoisie triumphed, and feudalism as an epoch gave way to capitalism. Unlike previous epochs, capitalism is based, not on agricultural production, but on industrial production. The dominant class, the *bourgeoisie*, created *bourgeois democracy* as a means to defeat the feudal lords and establish its supremacy. The other major class in capitalism is the *proletariat*. Members of the proletariat own none of the means of production, but they do own themselves. They sell their labor in exchange for wages. In the early period of capitalism, there were other classes, such as the *petite bourgeoisie* (little bourgeoisie), or those who owned some of the means of production (such as mom-and-pop merchants or owners of family farms). Over time, the petite bourgeoisie had been competed out of existence by larger, more efficient producers (the *industrial bourgeoisie*), and subsequently the petite bourgeoisie joined the ranks of the ever-expanding proletariat. Indeed, over time, a polarization of sorts would occur, with wealth being concentrated in fewer and fewer hands and the proletariat growing ever larger and more impoverished.

Somewhat surprisingly, Marx did not consider capitalism to be completely devoid of any positive impact on humanity. In fact, he claimed that capitalism provided a dynamic means to concentrate resources and convert those resources into unprecedented technical advances in very short order. Indeed, capitalism was the most efficient and productive epoch in human history. However, Marx asserted that this dynamism came at a severe price. First and foremost for Marx, capitalism facilitates an exploitative relationship between the two major social classes—the owners of capital (the bourgeoisie) and the working class (the proletariat). Briefly, Marx claimed that the profit (also known as *surplus value*) derived from the capitalist production process was merely the difference between the value generated by the proletariat and the wages that they earned from the bourgeoisie. Therefore, according to Marx's conception, the proletariat generated all value as a result of its labor but had only a portion of that value returned to it by the bourgeoisie in the form of wages. Since the proletariat created surplus value, but the bourgeoisie

enjoyed the fruits of that value, the bourgeoisie was effectively exploiting the proletariat on a consistent and ongoing basis.

Marx asserted that this exploitative relationship was an essential part of the capitalist production process. Among other things, surplus value was used by the bourgeoisie to reinvest, modernize, and expand its productive capacity. All members of the bourgeoisie had to expand the scope of their productive operations, or eventually they would be put out of business by rivals from within their own social class. Therefore, for Marx, capitalism could not continue as a mode of production without the unceasing exploitation of the proletariat, which comprises the majority of human beings in advanced industrial societies.

Not only did Marx claim that the capital–wage labor relationship was exploitative, but he also claimed that this economic relationship left the majority of human beings feeling estranged from their own humanity. Because Marx believed that productivity was a naturally human act, he concluded that the capital–wage labor relationship degraded something that was a fulfilling, meaningful, and free act into drudgery that was performed solely for the purpose of basic survival. Since humans constantly reproduced their material conditions and, in doing so, refashioned human nature, work performed for the sole purpose of survival ultimately served to alienate all members of the proletariat from their very humanity.

Marx predicted that capitalism, like every dominant economic mode of production before it, possessed internal contradictions that would eventually destroy the system. Not only was the everyday capital–wage labor relationship marked by exploitation, but the nature of the market system also guaranteed that the economy would slip into periodic crises that made the exploitative nature of the association between the bourgeoisie and proletariat clear for all to see. These would be revolutionary moments when the proletariat would achieve *revolutionary consciousness*, or the realization that the source of its misery was the system of capitalism itself, and would rise up and destroy it.

Classical economists of Marx's time recognized the negative impact that periodic economic recessions had on capitalist economies, but they generally viewed such downturns as acceptable (some even considered them positive events) and temporary. Marx, on the other hand, interpreted these recessionary periods as a sign of profound contradictions inherent in capitalism. These recessions were moments of crisis, Marx thought, and not necessarily temporary in nature. Furthermore, Marx predicted that, over time, crisis periods would get progressively longer, recessions would get deeper, recoveries would be shallower, and times in between moments of crisis would get shorter. Ultimately, like all other modes of production before it, Marx claimed, capitalism would come to an end and be replaced by an economic system that had fewer internal contradictions.

Socialism and Communism

Following the collapse of capitalism and the seizure of power by the proletariat, a transitional period would follow, *socialism*, ultimately leading to full-blown advanced communism. Marx spent very little space discussing his vision for socialism and communism, but he and Engels discussed it briefly in *The Communist Manifesto* (1848). He also referred to life under socialism in *The German Ideology* (1845), and he commented on the basic principles of socialism and communism in commentaries such as the *Critique of the Gotha Programme* (1875). From these indications, the following picture emerges. During the transitional period, the proletariat uses the coercive power of the state to defend the revolution from the remnants of the bourgeoisie. However, because the “habits” of the past are not easily discarded, Marx and Engels contended, some form of exchange would continue. In the *Critique of the Gotha Programme* (1875), Marx states that in a socialist society, the laborer will receive, in return for a given quantity of work, the equivalent in means of consumption, or the formula later adopted by the Soviet Union, “from each according to his ability, to each according to his labor.”

Over time, with continuing production comes the elimination of material want (a blessing of industrial capitalism that provided the productive basis for the communist epoch) as well as the disappearance of the last vestiges of the bourgeoisie. Under socialism, the proletariat would represent both the majority of society and the dominant class. Under communism, there would be no classes, because all would have equal access to the means of production. Production in such a system would be designed to serve human needs rather than extracting the highest possible levels of surplus value. Marx sums it up in the following words:

In a higher phase of communist society, after the enslaving subordination of the individual to the division of labor, and therewith also the antithesis between mental and physical labor, has vanished; after labor has become not only a means of life but life's prime want; after the productive forces have also increased with the all around development of the individual, and all the springs of co operative wealth flow more abundantly only then can the narrow horizon of bourgeois right be crossed in its entirety and society inscribe on its banners: From each according to his ability, to each according to his needs! (Marx, 1875, Part I, para. 50)

Full communism would have some key characteristics (Marx & Engels, 1848). It would be a classless society, because class differences would disappear. One might wonder why class differences would disappear, given that Marx's account of history was almost entirely based on class conflict as “naturally” arising. In part, their disappearance is one of the blessings of capitalist production. Capitalism is so materially productive that it would

produce such abundance that no one would want for anything. Technology had provided for such material abundance that there would be no need for haves and have-nots, that is, no classes. As a result, given that the state was seen as a tool of the dominant class, communism would ultimately be a stateless society as well, because the state would ultimately “wither away” of disuse. This idea was especially developed by Friedrich Engels in *Anti Duhring* (1877). Furthermore, communism would be a nationless society because, Marx and Engels believed, national identities were a product of capitalism, and such identities would disappear, to be replaced by a universalist proletarian identity. For Marx, under communism, people would be free to do all that they wish. He described life under communism in the following terms:

[Communist society would make] it possible for me to do one thing today and another tomorrow, to hunt in the morning, fish in the afternoon, rear cattle in the evening, criticise after dinner, just as I have a mind, without ever becoming hunter, fisherman, herdsman or critic. (Marx, 1845, para. 10)

Some argue that Marx defies his own foundational philosophy by declaring that socialism and communism are a historic inevitability. However, there is little evidence that Marx genuinely believed that socialism or communism represented the essential next step in human history. In fact, in the *The Communist Manifesto*, Marx and Engels confess that the conflict between the bourgeoisie and the proletariat may very well be vicious enough that it leads to the destruction of humankind rather than the dawn of a new millennial age.

In sum, Marx expends much more of his intellectual career analyzing and critiquing capitalism than he does defining the nature of socialism and communism or life after the revolution. In general, this squares with his philosophical roots. If history is defined by human activity and conflict, as Marx postulated, then it would be impossible for anyone to describe future modes of production in any level of detail—the details by necessity would be provided by those who refashion history.

The Role of the Professional Revolutionary

What then is the role of the professional revolutionary, if the laws of history appear to predetermine the inevitable collapse of capitalism? For Marx, although the proletariat has a historic mission, this mission is not always clear. Thus a revolutionary party is needed to enlighten the proletariat and help form it into a class, which would then lead to the overthrow of the bourgeois supremacy and the subsequent conquest of political power by the proletariat. The role of the professional revolutionary party was to prepare the proletariat for its revolutionary mission by educating the masses as to their historic purpose.

Marx certainly encouraged members of the proletariat to be conscious of their class status and organize as producers into revolutionary unions, political parties, and fraternal organizations. He would leave the specifics to the many who would follow in his wake.

Conclusion

Marxism has been one of the most influential political ideologies of the 19th and 20th centuries. Marx's ideas not only inspired a variety of schools of thought, but his ideas have inspired a vigorous debate over a whole range of issues—such as the balance of the state and the market in production and the proper role of government in society. Indeed, one of the main criticisms of Marx and Engels is their tendency to underestimate the power of the capitalist state to stave off the inevitability of revolution. Indeed, Marx did not foresee the power of the welfare state in saving capitalism from itself.

Furthermore, the number of schools of thought that have derived from Marx's ideas are too numerous to recite in a brief chapter such as this. However, those inspired by Marx fall roughly into two categories: *revolutionary socialists* and *evolutionary socialists*. Of all the revolutionary socialists, the writings of Vladimir Ilyich Ulyanov (better known as Lenin) stand out as the most prominent. The Western European tradition of social democracy, in which the interests of the proletariat are represented by a political party that seeks to gain power through democratic elections, offers a stark nonrevolutionary contrast to Leninism. In spite of the intense differences between these two schools of socialist thought, both undeniably owe their foundational ideas to the work of Karl Marx. They are subjects of other chapters in this handbook.

Although the study of Marxism after the collapse of the Soviet Union in 1991 has gone out of vogue in many intellectual circles, its relevance now has become increasingly apparent. The concentration of wealth in fewer and fewer hands via corporate mergers and hostile

takeovers, the disappearance of the petite bourgeoisie (family farmers and mom-and-pop enterprises), and the apparent collusion between big capital and the state—all were suggested by Marx and Engels. Perhaps a rediscovery of Marxism among students of political science would help them better understand the direction of the world in the 21st century.

References and Further Readings

- Barnett, V. (2009). *Marx*. London: Routledge.
- Engels, F. (1877). *Anti Duhring*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1877/anti-duhring>
- Engels, F. (1884). *The origin of the family, private property and the state*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1884/origin-family/index.htm>
- Evans, M. (1975). *Karl Marx*. London: Macmillan.
- Marx, K. (1844). *A contribution to the critique of Hegel's philosophy of right*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1843/critique-hpr/intro.htm>
- Marx, K. (1845). *The German ideology*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1845/german-ideology>
- Marx, K. (1867). *Capital*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1867/c1>
- Marx, K. (1875). *Critique of the Gotha programme*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1875/gotha/index.htm>
- Marx, K., & Engels, F. (1848). *The communist manifesto*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1848/communist-manifesto>
- Marx, K., & Engels, F. (1884). *The Eighteenth Brumaire of Louis Napoleon*. Retrieved April 5, 2010, from <http://www.marxists.org/archive/marx/works/1852/18th-brumaire>
- McLellan, D. (1973). *Karl Marx: His life and thought*. New York: Harper & Row.
- Mehring, F. (2003). *Karl Marx: The story of his life*. London: Routledge.
- Tucker, R. C. (1978). *The Marx Engels reader*. New York: Norton.
- Whean, F. (2002). *Karl Marx: A life*. New York: Norton.

77

REVISIONISM AND SOCIAL DEMOCRACY

JUTTA A. HELM

Western Illinois University

Even by contemporary standards, the 19th century witnessed dramatic changes in Europe's political, social, and cultural life. Revolutionary upheavals, the slow advance of democracy, scientific breakthroughs, and new ideologies challenged the status quo. But the most dramatic change resulted from the steady advance of capitalism in what Polanyi (1944/1962) called *The Great Transformation*. What was most revolutionary about capitalism was the creation of a market for human labor that subjected workers and their livelihood to the law of supply and demand. Stripped of the protections that traditional communities, with their networks of obligations and duties, had provided, workers were now compelled to become wage laborers in factories, mines, and farms, where harsh working conditions, low pay, and frequent unemployment clashed with capitalism's promise of progress and prosperity.

Utopian socialism, with its emphasis on public ownership of economic resources and an egalitarian vision of society, was one response to capitalism. Robert Owen in Great Britain and Charles Fourier and Pierre-Josef Proudhon in France, among others, invented various socialist theories in the first part of the 19th century. Some even tested their theories in communal experiments in the United States, Great Britain, and France. In 1848, a new brand of socialism burst on the scene with the publication of *The Communist Manifesto* in Paris in 1848. Its authors, Karl Marx and Friedrich Engels, claimed to offer workers

a scientific theory of socialism that promised to liberate workers and lead them toward the eventual creation of a communist society based on the ideals of justice and equality. These ideas were articulated more fully in their later writings and in their political work with trade unions and working-class political parties all over Europe. Initially, these parties had espoused various socialist viewpoints. Most called themselves *socialist* or *social democratic*, labels they used pretty much interchangeably. But by the 1890s, most of them had committed to the Marxist version of socialism, and by 1914, virtually every European country had at least one Socialist or Social Democratic party.

The widespread adoption of Marxist socialism by labor unions and political parties is due in large part to the important role that the German Social Democratic Party (SDP) played in Europe's working-class movement. With close to 20% of the vote in 1890, it was the largest party in the Second Socialist International, a loose alliance of more than 20 labor and working class parties that was active from 1889 to 1916 throughout Europe. The SDP's leaders, August Bebel and Karl Kautsky, did much to adapt and popularize Marxist ideas for a larger audience. In fact, by the 1890s, their interpretations of Marxism were read more widely than Marx's own works, and they were translated into many languages. They also closely collaborated with Friedrich Engels, who had become the guardian of Marxism after

Marx's death in 1883. So it is ironic that the major challenge to the Marxist theory of socialism—*Revisionism*—came from the inner circle of the very party that had done so much to promulgate Marxism in Germany and in Europe.

Democratic Revisionism

Eduard Bernstein, a largely self-educated Marxist intellectual, contributed the most to the development of revisionism. It stands for a vision of socialism that revises and partially rejects Marxism in favor of a more gradual movement toward socialism embedded in democratic values and institutions. In 1872, Bernstein became an eager convert to Marxism. He quickly won the respect of party leaders, and in 1878 he joined many of them in exile, first in Switzerland and then in London. The German government had passed the so-called Anti-Socialist Laws banning socialist meetings, publications, and all organizing activities for the next 12 years. Bernstein was chosen as the editor of *Der Sozialdemokrat*, a monthly journal serving as the official publication of the SDP in exile.

During his time in Switzerland (1878–1888) and in London (1888–1900), Bernstein attended all the important underground meetings of the party, and he met and worked with the most brilliant and dedicated leaders of German and European social democratic parties. Among these was Karl Kautsky, an Austrian social democrat who established himself as a skillful ideologue who worked closely with Bernstein and August Bebel, the leader of the SDP. Bernstein published many articles that drew on the German government's repression of political opponents and the continuing economic depression (1873–1896) as evidence of the validity of Marx's predictions about the deepening crisis of capitalism and its eventual downfall.

In 1888, the Swiss authorities forced Bernstein and several other Social Democrats to leave, and at Engels's urging, Bernstein moved to London. Subsequently their relationship grew very close, leading Engels to confide in a friend that in matters of strategy and theory, he trusted Bernstein as much as himself. Eventually he chose Bernstein and Bebel as the executors of his will. In 1890, Bernstein's position within the SDP underwent considerable change. Due to the expiration of the Anti-Socialist Laws in Germany, there was no longer any rationale for the publication of *Der Sozialdemokrat*, given its purpose of maintaining the party's intellectual life in exile. Instead, the party offered Bernstein the position of London correspondent for two other publications, *Neue Zeit* (New Time), edited by Kautsky, and *Vorwaerts* (Onwards), based in Berlin. To Bernstein's delight, these assignments proved much less time-consuming than his previous job. However, he lost an influential platform for influencing intraparty debates. This became important in 1891, when, boosted by its impressive showing in the 1890 elections, the SDP prepared the draft for a new party program to be

debated and adopted at the party conference in Erfurt. Known as the *Erfurt Program* (n.d.), the document included Kautsky's theoretical part, hewing closely to Marxist principles, and a practical part, written by Engels and Bernstein, summarizing an action program calling for democratic reforms and improved labor laws.

To Bernstein, the two parts of the program were entirely unrelated: Although the theoretical section focused on the inevitable collapse of capitalism, consistent with Marxism's claim of a scientific analysis of the present and future, the practical part listed reforms such as universal suffrage, the right to free expression, effective worker protection laws, and legal equality for women, to mention only a few. Many non-Marxist reformers supported these proposals. Oddly, the abolition of private ownership of economic resources—the core feature of Marxism—was not included. The program perpetuated the sharp division of earlier programs between revolutionary theory and reformist practice. If the goals of the working class could be advanced by reform, was it necessary or even desirable to call for the abolition of capitalism? Was it necessary to “modernize” Marxist theory? These questions increasingly troubled Bernstein.

In London, Bernstein maintained a busy schedule as a writer, journalist, and publisher. His wide-ranging social and political contacts included labor leaders, intellectuals such as Beatrice and Sidney Webb (leaders of the British Fabian Society, which advocated a gradual move toward socialism), artists such as William Morris, Christian Socialists, and Left Liberals, among others. Bernstein became a regular contributor to current affairs journals such as *The Nation* and *Progressive Review*. In many ways, he felt quite at home in London, enjoying the long tradition of free speech that allowed for lively public debates without fears of censorship, as in Germany. But he also longed to return to Germany to influence debates within the SDP. He felt these debates were essentially frozen as the party leadership continued to cling to the three pillars of orthodox Marxism it had embraced earlier. What were these pillars?

First, there was the Marxist belief in the *collapse theory*. As Marx and Engels (1848) wrote in the *Manifesto*, deeper economic crises will eventually lead to revolution as the workers rise up against the system that enslaves them. Or, more eloquently, “What the bourgeoisie, therefore, produces, above all, are its own gravediggers” (Marx & Engels, 1848, chap. 1, para. 53). Hence, the collapse of capitalism is inevitable. The second pillar of communist orthodoxy is the *immiseration theory*. Accordingly, as capitalism develops, workers are increasingly reduced to appendages of the machines they operate, living only so long as they have work, and working only as long as they can increase capital, or profit (Marx & Engels, 1848). As more and more workers become paupers, they are joined by impoverished craftsmen, even the middle class, to the point that there are only two classes: the ever more desperate working class, or *proletariat*, and the ever more

powerful capitalist class, or *bourgeoisie*. Third, Marxists emphasize *historical materialism*, the notion that social and political developments are determined by economic forces. Far from being just an idle philosophical quarrel, the insistence on historical materialism suggests that political reform is of limited usefulness. Instead, the working class needs only to wait for capitalism to reach its final crisis, at which point workers will rebel as they “have nothing to lose but their chains” (Marx & Engels, 1848, chap. 4, para. 11).

Bernstein was increasingly troubled by these tenets. His critique focused on practical difficulties as well as philosophical inconsistencies. The practical difficulties loomed larger and larger as the progress of democracy in Germany and other European countries presented the representatives of working-class parties with several dilemmas. Should they take advantage of the possibility of improving working-class lives through legislation, or was parliamentary work just a diversion from the revolutionary struggle? Would reforms blunt the wave of the coming revolution? In Germany, a group of Social Democratic legislators called the *practitioners* advocated that tangible improvements for workers should be pursued, a view not shared by the leadership. Should Social Democratic legislators seek to build alliances with nonsocialist parties to advance the cause of reform? Here too the party insisted that such alliances were not desirable unless the Social Democrats were in control, a situation that was precluded by their minority status. In several countries (France and Italy, among others), Socialists had been invited to join nonsocialist governments. Should they accept the invitation for the sake of opening up more avenues for meaningful reforms? This question deeply divided Socialist parties, to the point that by World War I, not a single Socialist legislator had, with the approval of his party, participated in a government.

On a philosophical level, Bernstein doubted the Marxist claim to have discovered a scientific explanation of past and future history. He took issue with this claim when he pointed to several discrepancies between Marxist predictions and current reality. Marxism, he insisted, was not a rigid doctrine set in stone by its authors. It is the duty of their followers, he argued, to remove contradictions and to develop it further. In this sense, revisionism was a call for fresh thinking as opposed to “everlastingly repeating the words of [the] masters” (Bernstein, 1911, p. 26). So in 1896, Bernstein published the first of several articles in which he challenged what he saw as rigid dogma and offered “revisions” to put social democratic and socialist theory in Germany and Europe on a firmer footing. Three issues were paramount in these articles. First, Bernstein (1896/1988a) urged the parties to begin discussions of what socialist and communist societies were to be like. Instead of assuming “an abrupt leap from capitalist to socialist society” or “a decisive victory of socialism,” there should be serious thinking about the transformation (p. 74). How was it going to come about? Was revolution the only avenue

toward socialism? Or were there other avenues leading through democratic governance in which a stronger socialist party would create the conditions for socialism?

Second, Bernstein (1898/1988b) argued that it was increasingly unrealistic to expect the inevitable collapse of capitalism. Instead, he suggested that capitalism was resourceful and capable of adjusting to avoid its self-destructive tendencies. This was true even in Germany, he argued, where the end of the long depression of 1873 showed an invigorated economy that created new wealth among broader groups of society. Of course there would be future crises, but most likely they would be contained. In his third assault on old dogma, Bernstein rejected the immiseration theory. Based on his analysis of Prussian census data, he argued that society was becoming more, not less, differentiated into various middle-class groups. The peasantry too was holding its own. This argument had important practical implications for Bernstein. Since the nonproletarian groups in society were not about to disappear, it might become useful and even necessary to look to them as potential allies for reformist causes. Taken together, these three criticisms form the core of revisionist thinking.

Published just prior to the SDP’s party conference in Stuttgart in 1898, Bernstein’s critique ignited a furious debate. Some charged him with creating a full-blown crisis for the party. Others accused him of having caught the “British disease” of embracing gradualism over revolution. Others still, notably the Left Socialist faction and Rosa Luxemburg, chided him for his wholesale betrayal of socialism and called for his expulsion from the party. But most delegates were content to close ranks and affirm the ideological unity of the party and its Marxist dogma, an outcome that was carefully orchestrated by the party leader, Bebel, and Kautsky, Bebel’s chief ideologue. There can be little doubt that their leadership positions would have been in jeopardy had Bernstein met with greater support at the conference.

Some time later, Bernstein conceded that his challenge of party dogma was perhaps asking for too much too quickly. But he never wavered. Before returning to Germany in 1900, he wrote his revisionist manifesto, *The Preconditions of Socialism*, or, in its English translation, *Evolutionary Socialism* (1911), which presented his ideas in a systematic exposition. As socialist works go, it is relatively brief (about 200 pages), as Bernstein wanted it to be accessible to workers as well as party activists. It is important to note, however, that Bernstein’s thinking continued to evolve through the 1920s. He came to disagree with his old comrades about the role of democracy as a mere facilitator of socialism. Democracy, he argued instead, provided the means for a radical transformation of society through the principled and persistent struggle for a more just, equitable world. But democracy was also an end because of its focus on politics as the arena for change. Here socialists could join with other groups and parties, motivated by their common humanity and a vision of a

better world, using the power of the democratic state to reshape the world around them (Berman, 2006). This view was strongly rejected by the party elite, which viewed democracy as only a stepping stone, a mere way station toward socialism.

Although never repudiating Marx, Bernstein worked hard to show how his evolutionary-liberal vision of socialism was compatible with Marx. They shared an optimism about a better world without exploitation. But could the steady reformism advocated by Bernstein achieve a radical transformation of capitalist society? This question retained its urgency as later generations of revisionists tried to implement his program. Bernstein was able to pursue his commitment to joining theory and political practice as a deputy in the German parliament from 1902 to 1928. He worked tirelessly to refine his ideas and probe practical, policy-oriented issues such as a more just tax policy, trade policy, and constitutional law, to mention just a few. He also enjoyed considerable public support, especially among trade unions, even though many party leaders remained openly hostile. And his influence spread well beyond Germany. *The Preconditions of Socialism* was translated into more than 20 languages, and soon there were revisionist factions in most socialist parties. As Lenin, the leader of the Russian Social Democratic Party and an astute observer of developments in Germany and Europe, noted in 1901, “The French socialists have begun, not to theorize, but to act. The democratically more highly developed political conditions in France have permitted them to put ‘Bernsteinism into practice’ immediately, with all its consequences” (Lenin, 1901, chap. 1, para. 5).

Nonetheless, revisionism remained controversial, and none of the major socialist parties formally embraced it before World War I. Several parties, such as those in Italy and France, split over the issue, with leaders such as Jean Jaures in France and Francesco Merlino in Italy heading revisionist parties.

With the onset of war in 1914, Europe descended into considerable social and political turmoil, and socialist unity across nations collapsed as nationalism came to the fore. And even though socialists joined patriotic unity governments in a number of countries, it seems that they were completely unprepared for the war and its aftermath. The SDP in Germany, for example, enjoyed considerable power and electoral support, but there was no socialist breakthrough once the party found itself in the government. Blockage and impasses frustrated the party at every turn. Revisionism remained a contested concept, and the socialist movement was now fractured into the new communist parties that advanced the idea of revolution on the Russian model, traditional socialists clinging to orthodox Marxism in their rhetoric but exceedingly cautious in their actions, revisionists adhering to Bernstein, and pragmatists consumed by the day-to-day struggle without any coherent vision about the future. Nationalist parties emerged as new competitors, and most socialists were unable to develop a viable response. Bernstein and several French and Austrian

leaders had argued that workers had national and international commitments, challenging the notion that nationalism was merely a tool of the ruling classes, but this idea was rejected by the orthodox groups.

The first major challenge for socialist parties came in Italy, where the democratic order was threatened by *fascism* in 1923. Italy’s Socialist parties were not able to muster effective countermeasures. The second major challenge came in 1932 and 1933 in Germany, where the Great Depression that began in 1929 had had devastating economic effects. The national socialist movement, led by Adolf Hitler, and the Communist Party benefitted from widespread public despair. There were revisionist responses to the Great Depression, but they were not implemented. In Germany, the so-called WTB Plan, named after its sponsors, called for 2 trillion deutsche marks to be spent as an economic stimulus for work creation. Trade unions were desperate to give workers some hope and stem their defection to nationalist and communist groups. But the SDP could not bring itself to accept and implement this plan because, in the words of its major economic spokesman, Rudolf Hilferding, it was “not Marxist” (cited in Berman, 2006, p. 114). Ironically, Hitler’s government, coming to power only a few months later, did implement such a plan, with quick results. Another plan, called *Plan du Travail* (Plan of Work), was proposed by the Belgian Socialist Hendrik de Man, who advocated short-term work creation projects to fight the Depression and a long-term plan to reshape capitalism (Berman, 2006). Going one step further than Bernstein in dropping Marxism completely, de Man advocated that class solidarity give way to social solidarity in a reformist socialism that focused on control of the means of production, not public ownership. The struggle was not against capitalism, he argued, but against particular types of hypercapitalism that exploited workers. De Man’s plan did have some impact in Belgium, Switzerland, the Netherlands, and France, where socialists seem to have understood more clearly that inaction was likely to benefit the radical right and the communists on the left. But it was in Sweden and the Nordic countries that revisionism had its greatest success. Here a tradition of political reform, in combination with effective leadership, provided the seed bed for the implementation of revisionist ideas on a national scale.

From Revisionism to Social Democracy

Sweden’s socialist party was relatively young. Soon after its founding in 1889, the *Socialdemokratiska Arbetarepartiet* (SAP) had embraced revisionism. Marxism was seen as a guide, not a dictate. The party valued democracy as an end in and of itself, and it was not averse to cooperation with nonsocialist parties. Unlike the SDP in Germany, the SAP saw its share of the vote in national elections rise steadily during the 1920s, when centrist and conservative governments were in power. The party honed its message to broaden its

appeal, searching for a “third way” between orthodox Marxism and exploitative capitalism. Just before the Great Depression, the party leader, Per Albin Hansson, coined the concept of “the people’s home” to convey the SAP’s desire to transform Sweden into a society where citizens cooperate on the basis of equality and helpfulness, without the economic barriers that separate the privileged from the neglected groups (Tilton, 1990). The aim was not to eliminate capitalism but to transform it through incremental reforms. Nationalization of industries played a rather small role in the SAP’s vision. The most crucial part of its promise was the effort to counteract the massive unemployment created by the Great Depression. A small group of party leaders decided to promote employment creation through government spending, not unlike the plans that were discussed in Germany and in Belgium and, a little later, in Great Britain. But the Swedish plan was homegrown, and it seems to have bolstered the party’s appeal. In 1932, after 3 years of economic turmoil, it garnered almost 42% of the vote. A year later, it forged a successful worker–peasant alliance with the Agrarian Party that gave the government a firm base of support for the implementation of its program: employment creation programs financed through government investments and welfare state policies such as housing assistance, aid to large families, paid holidays, and indexed pensions, to mention the most important ones. By 1936, unemployment had come down considerably. This created the basis for an agreement between labor and business that gave both sides what they wanted: Labor achieved collective bargaining rights, and business got a code of regulations for the management of industrial relations. This agreement, called the *Saltsjobadan Agreement*, effectively regulated labor–management relations for the next 40 years.

By 1940, Swedish socialists no longer talked about class conflict or nationalizations. They had learned to use the power of government to tame capitalism. This success was re-created in Norway and Denmark. They succeeded in showing that governments can bring order to chaotic economic markets, with benefits that are broadly shared. They took revisionism further than others, and in the process they laid the foundation for what would become the West European idea of social democracy after World War II.

Social Democracy Defined

Like revisionism, theories of social democracy developed out of the practical experience of West European socialists seeking to correct the shortcomings of capitalism and democracy and promote a more just society. Practitioners like Hansson from Sweden and scholars like Britain’s T. H. Marshall (1950) focused on the idea of citizenship to develop a new paradigm for a democratic society. Citizenship, they argued, is three-dimensional: There are (1) civil rights such as legal equality, (2) political rights, and (3) social rights. The latter confer economic security and resources that empower citizens to benefit from the

opportunities privileged groups have enjoyed all along. Put differently, without social rights, the promise of democracy remains hollow. Beiner (2001) describes social democracy as a uniquely European response to the question of what full democracy requires. What is needed, according to social democratic theory, are three crucial components: economic policies that regulate capitalism, social welfare programs that moderate inequality, and democratic governance. In addition, there is agreement that social democracy cannot be defined as a fixed set of rules and requirements. As capitalism changes, so must social democracy. Based on the Swedish experience of the 1930s and 1940s, social democracy must adjust to changing circumstances without abandoning its core principles.

At the end of World War II, Europeans had little faith in the regenerative strength of capitalism. In fact, fresh memories of the Great Depression and of wartime deprivation made capitalism outright unpopular. Communism was even more unpopular, and parties that called themselves socialist or social democratic took pains to distinguish themselves from the ruling communist parties in the Soviet Bloc countries behind the Iron Curtain. Communist parties in Western Europe were largely excluded from any participation in governments. Socialist or social democratic parties did enjoy considerable support, with at least a third of the electorate behind them. Only in Sweden, Norway, and Great Britain were they able to govern alone; elsewhere they joined coalition governments with other parties. Business groups were politically weak, especially in countries like Germany, Austria, and France, where they had collaborated with National Socialist governments. This finally opened the door to the kinds of reforms that had a distinctly social democratic flavor.

Developing the Welfare State

In the immediate aftermath of World War II, the need for policies to improve living standards was overwhelming. Most countries already had a rudimentary system of social policies in place—namely, old age pensions, health insurance, and occupational injury insurance. But now these policies were expanded. The National Health Service in Great Britain, beginning in 1947 and which created a universal health care system, is a good example of the social democratic approach to the welfare state. It is a noncontributory welfare program: All citizens are entitled to free care financed out of general tax revenue, regardless of their earnings or personal wealth. This entails some redistribution of resources as those with higher earnings and paying higher taxes contribute more than those with lower or no earnings (i.e., children). Not surprisingly, such policies were most widely adopted in countries where Social Democrats enjoyed a long period of political power, notably in Scandinavia, Austria, Belgium, and the Netherlands. A strong, unified trade union movement also helped boost welfare programs. In most countries, new welfare-state policies were less

generous and less comprehensive. But even conservative governments could not ignore the widespread demands for more benefits, and most responded, occasionally stealing the socialists' thunder. But in these countries, benefits were more likely to be means tested rather than universal, or, with pensions, for example, benefits were closely tied to earnings. The most important benefits include public pension systems, national health care programs, unemployment and disability insurance, housing subsidies, maternity and child benefits, and parental leave programs. While there was some retrenchment in welfare spending during the 1980s and 1990s, social democrats and trade unions resisted more drastic cuts. Today social spending amounts to almost 30% of gross domestic product in Sweden and France, with average levels of 24% for the 27 countries belonging to the European Union. In contrast, social spending in the United States amounts to about 15% of gross domestic product (Organisation for Economic Co-operation and Development, 2010).

Reforming Capitalism

An expanded welfare state requires economic growth. What ideas did socialists and social democrats have for the reform of capitalism to reduce its inherent inefficiencies and inequitable outcomes? Here historians of the postwar era see little systematic thinking. By and large, parties on the left side of the political spectrum advocated piecemeal reforms to improve capitalism. One such reform was the nationalization of key industries. This was quite popular in the immediate postwar years. In France, many leaders on the left and right advocated nationalization and economic planning as tools to resurrect the country after the devastation of World War II. Utilities, transportation, and a few banks and telecommunications were nationalized and continue in public ownership today. In Great Britain, nationalizations were more extensive and included railways, coal mining, utilities, and the iron and steel industry. These industries were reprivatized during the 1980s. Social democrats and some conservative parties in Germany and Austria proposed nationalization of basic industries, but in Germany at least, this was vetoed by the United States, which was one of the four powers that occupied Germany until 1949. Today there is a good deal of skepticism about the effect of nationalizations. Did they contribute to economic growth or full employment? As the nationalized industries were required to be commercially profitable, it is uncertain whether private management would have been much different. This is probably one of the reasons that Scandinavian social democrats did not nationalize any industries.

A German innovation did have a distinctive social democratic theme, however. Between 1951 and 1955, trade unions were able to muster enough strength to persuade the conservative government to pass three laws establishing a form of industrial democracy, or *codetermination*. The most radical of these laws applies to the coal and steel

industries—dominated by companies that had closely collaborated with Hitler's policies. Here, elected employee representatives sit on the companies' supervisory boards, and the board of directors includes a labor director whose selection requires union approval. Two other laws extended a watered-down participatory component to other large firms, where employees elect a works council that can negotiate a number of issues (excluding wages) with management. This law was revised in the 1970s to strengthen employee participation rights. Although the laws fall short of establishing genuine industrial democracy, in which employees and managers meet on an equal footing, they are generally given credit for ushering in a long period of social peace or *social partnership*, a term that remains very popular in Germany (Helm, 1986). Codetermination also supports the ideal of social citizenship, giving employees a say, albeit limited, in the management of businesses. Modified versions of codetermination have been implemented in a number of other European countries.

By the 1960s, European economies had stabilized, and social democratic debates about restructuring capitalism seemed less urgent. There were new labels for Europe's postwar capitalism, such as the "mixed economy," or the "social market economy," but they mostly described old wine in new bottles. A number of economies were close to achieving full employment, meaning the jobless rate did not exceed 3% of the labor force. Wages were rising, and workers enjoyed greater social protection through expanded social benefits. Many social democrats became "realists" and focused on maintaining their electoral appeal. They transformed themselves from working-class parties to "people's parties" in order to more effectively reach out to white-collar voters. The emphasis in party programs shifted from reforming capitalism to the promise of doing a better job at ensuring growth and full employment than conservative or centrist governments would. In Great Britain, Anthony Crosland's (1956) book, *The Future of Socialism*, reflected this new approach. Socialism, he argued, comes gradually. In order to deliver equal opportunities and more social spending, governments require economic growth, which can be sustained with the help of the countercyclical economic policies proposed by John Maynard Keynes (1936/2007) in 1936. This change to realism was most visible when the German SDP debated its new *Godesberg Program* in 1959. All references to Marx, class conflict, and the collapse theory were dropped. Socialism was now defined in ethical terms, inspired by traditions ranging from Christian ethics and humanism to classical philosophy. Bernstein's vision finally triumphed in the party he had worked so hard to reform! The new program retained the optimism and belief in progress that Bernstein shared with Marx. The economic climate of the 1960s seemed to validate this optimism. Some referred to this decade as "the golden age of capitalism," with double-digit rates of economic growth, rising incomes, and higher social spending.

But the next decade changed all this. With dramatic rises in oil prices, inflation, and unemployment, social democratic goals seemed to become elusive again. Whereas during the 1960s the conflict with conservatives had been about the distribution of the economic surplus, now questions about government's role in the economy rose to the top of the political agenda. Conservatives such as Britain's Margaret Thatcher demanded lower taxes, less regulation, and lower social spending while social democrats argued for an expanded regulatory regime to curtail "excessive" wage growth. There was widespread concern that higher unemployment was due to high wages. The Swedish experience seemed telling. Here, centralized wage bargaining since the 1950s had put pressure on companies to achieve high rates of profit, parts of which went to satisfy union demands for higher wages. This system, called the *Rehn Meidner model*, also had an equalizing component in that it was designed to reduce wage differentials between skilled and unskilled workers (Sasson, 1995). Beginning in the late 1960s, skilled workers were less willing to accept this solidarity-oriented system. Swedish governments now were in the awkward position of having to persuade unions to accept smaller wage increases. In Germany, the United Kingdom, Italy, the Netherlands, and Austria, governments faced the same dilemma. The new watchwords were *incomes policy* or *concerted action*, with social democratic governments using their political capital to persuade unions and employers to moderate wage demands and price increases. This new approach worked, at least in the short term, but sooner or later, unions, business groups, or both rebelled against its regulatory constraints. More often, business groups bolted, attracted by conservative and centrist parties that promised to reduce regulation in order to unleash market forces.

In the 1980s, social democrats found themselves and their ideas in a real crisis. The welfare state was under pressure from conservative governments eager to reduce taxes and government spending. Full employment seemed to be a thing of the past, and trade union rights were curtailed in a number of countries. The social democratic share of the vote in national elections remained at about the same level as in the 1950s. There were significant losses only in the United Kingdom and Germany, but there were also gains in France, where the first leftist government took office in 1981. Social Democrats gained power in Spain in 1982, the first government of the left since the 1930s. Even so, social democrats lacked confidence in the future. The Socialist International started the postwar era in 1951 with the promise to abolish capitalism. In 1989, its slogans focused instead on freedom, social justice, and solidarity. There was a general acknowledgment that social democracy would have to be reinvented.

Expanding Democracy

Beginning with Bernstein, social democrats had come to embrace democracy as an end in itself. In the post-World

War II era, they set out to change formally democratic societies into societies enjoying real democracy, bolstered by economic security and wider opportunities for all. When Willy Brandt, the leader of the German SDP, campaigned in 1969, he called for a government that would "risk more democracy," a slogan that had broad resonance among younger voters and intellectuals. Few were certain what this might entail. But the next two decades brought social democrats in Germany and elsewhere face to face with a new antiestablishment culture. Anthony Crosland (1956) had warned the left 15 years earlier that sooner or later, it would have to deal with totally new cultural issues such as concerns for civil liberties, personal lives, and leisure activities. In the 1970s and 1980s, this came to pass in several areas, of which this chapter discusses three. First, in a number of European countries, educational opportunities had remained relatively unchanged. Children from working-class families found themselves, not excluded from, but severely underrepresented in, institutions of higher education. Social democrats promoted primary and secondary schools that would be less reliant on early selection into different educational channels. In Great Britain, the notorious "11-plus exams" came to be seen as a barrier to talented working-class children whose development did not quite fit into the prescribed schedule. In Germany, access to the academic high schools was liberalized during the 1970s. Comprehensive schools, more years of required schooling, and greater access to universities and higher education were, as Sasson (1995) put it, "the standard left-wing position on education" (p. 393), and it became a common feature of Europe's educational systems. Similarly, social democrats were more interested than conservatives in expanding university access for women and, later in the 1990s, for immigrant children.

The second cultural issue took social democrats by surprise: the rise of the women's movement. Socialism had a long tradition of calling for sexual equality. Marx and Engels had argued against traditional family structures that "enslaved" women. In 1891, the *Erfurt Program* endorsed women's right to vote. But initially, social democrats were taken aback by the radicalism of some new feminist demands, especially with regard to sexual mores. The legalization of abortion became a central issue, and starting with Great Britain, legalization eventually reached the southern European democracies, where Catholic institutions were still strong. After legislation supported by Italian socialists and communists, Italian voters ratified legalization in a 1976 referendum. Economic issues were also on the feminist agenda. Unequal pay, occupational segregation, and lack of child care were now debated and eventually addressed, more or less adequately, by parliaments, employers, and courts. Equally interesting is the gradual arrival of more and more women in politics. In 1975, only the Scandinavian countries could boast double-digit representation of women in parliaments. Today, the European average stands at 21%, and in the Scandinavian

countries it is 41% (Inter-Parliamentary Union, 2010). Some social democratic parties adopted goals or even quotas to increase the number of female candidates for office. It is interesting to note that many conservative parties, after initially hostile remarks about the recruitment of “unqualified” women, felt obliged to follow suit, usually with less stringent requirements. The really important question is how the arrival of a critical mass of female legislators affected the business of legislating and governing, but this is addressed in another chapter.

The environmental movement posed the greatest challenge to social democratic parties. Given the fact that environmentalists usually call for more stringent regulation of industry, one might assume that social democrats had little trouble integrating them into their political fold. But this was definitely not the case. While all social democratic parties eventually wrote environmental party platforms, this may have been too little, too late. Many green movements were ideologically divided between “realists and fundamentalists,” between rural romantics, futuristic visionaries, and other factions. In the 1970s, green parties emerged in most European countries, and now they are represented in most parliaments. They compete for votes with social democrats but also cooperate with them once elected. The best known example is probably the social democratic–green coalition that governed Germany from 1998 to 2005. Like feminism, environmentalism forced social democrats to question their tradition in order to open up to new issues outside of their traditional ideological terrain.

Future Directions

Today, social democrats can claim credit for an impressive list of achievements. Writing in 1993, Przeworski (2001) argued that “the only countries in which almost no one is poor after taxes and transfers are those that pursue social democratic policies” (p. 778). A few years later, Sasson (1995) concurred with this assessment, adding that those countries also championed civil rights and democracy, fighting for the expansion of the right to vote where it was restricted, for the rights of women and homosexuals, and for the abolition of racial discrimination and capital punishment. And the social democratic welfare state may have saved European capitalism during its worst crises in the 1980s and over the past few years. Does it offer the same hope for the present economic challenges, or has it exhausted its usefulness? David Marquand’s (1993) cautionary remarks are even more appropriate for our times:

The capitalist free market is a marvelous servant but a disastrous master. In one of the greatest achievements of the second half of this [20th] century, a few favored societies learned to convert it from master to servant. The danger now is that a smug and vain glorious capitalism will not remember the lesson. (p. 51)

What, if anything, can social democracy contribute to the solution of capitalism’s current problems? There are at least two responses to this question. One view, taken mostly by neoliberal authors, is that the social democratic era is over. Many go further and argue that social democracy has saddled European welfare states with high taxes, sluggish growth, and low levels of innovation and productivity (Steyn, 2009). These critics argue for the burial of the social democratic model of society. Others, notably economic historians and policy analysts, argue that this model is as relevant as ever (Jacobs, Kent, & Watkins, 2003; Judt, 2009; Krugman, 2010). In their view, the social democratic challenge today is threefold. First, it is important to conserve the achievements of the 20th century. When neoliberals call for flexible labor markets, lower taxes, and fewer regulations, social democrats must rally to defend the reforms that humanized capitalism through policies that protect workers with a social and economic safety net. In the context of global capitalism, this is no small task. And it is in the global arena that social democrats confront the second challenge.

As Jacobs et al. (2003) argue, if social democrats want to help shape the capitalism of the 21st century, they must move their struggle to the global stage. Here international corporations have already established a foothold, taking advantage of production sites that offer low taxes, little regulation, and desperate workers accepting jobs without the protections social democrats have achieved. For this reason, social democrats have often demonized globalization and institutions such as the World Bank, the International Monetary Fund, and the World Trade Organization as indifferent or outright hostile to the concerns of working people. The task now is to democratize these institutions and to strengthen their weak regulatory powers to respond to the interests of workers as well as business. This is a huge challenge, and so far only tentative steps have been taken. There are some modest success stories at the level of the European Union, where the adoption of the Social Charter in 1992 established a regulatory structure that gives working people a stronger voice.

The third challenge for social democracy is a moral one. When much of our political discourse is dominated by *economism*, a single-minded concern with profit, productivity, and growth, a moral critique of the status quo often seems “soft,” well-intentioned but unrealistic. But as all too many people around the globe are experiencing growing inequality, insecurity, and fear, social democracy has an opportunity to confront these fears, to explore options and offer solutions that address them. There is now a global constituency looking for a message of hope, a promise that something can be done about the darker side of global capitalism, that growing inequality and insecurity are not inevitable (Judt, 2009). Can social democracy define new approaches to these

problems, approaches that create a new balance between capitalism and democracy?

References and Further Readings

- Beiner, R. (2001). Democracy, social. In S. M. Lipset (Ed.), *Political philosophy* (pp. 359–364). Washington, DC: Congressional Quarterly Press.
- Berman, S. (2006). *The primacy of politics*. New York: Cambridge University Press.
- Bernstein, E. (1911). *Evolutionary socialism*. New York: B. W. Huebsch.
- Bernstein, E. (1988a). General observations on utopianism and eclecticism. In H. Tudor & J. M. Tudor (Eds.), *Marxism and social democracy* (pp. 73–81). New York: Cambridge University Press (Original work published 1896)
- Bernstein, E. (1988b). The struggle of social democracy and the social revolution: The theory of collapse and colonial policy. In H. Tudor & J. M. Tudor (Eds.), *Marxism and social democracy* (pp. 159–173). New York: Cambridge University Press. (Original work published 1898)
- Crosland, A. (1956). *The future of socialism*. London: Jonathan Cape.
- Erfurt Program. (n.d.). Retrieved May 28, 2010, from [http://www.marxists.org/history/international/social democracy/1891/erfurt program.htm](http://www.marxists.org/history/international/social%20democracy/1891/erfurt%20program.htm)
- Gamble, A., & Wright, T. (1999). *The new social democracy*. Oxford, UK: Blackwell.
- Helm, J. A. (1986). Codetermination in West Germany: What difference has it made? *West European Politics*, 9(1), 32–53.
- Inter Parliamentary Union. (2010). *Women in national parliaments*. Retrieved May 25, 2010, from [http://www.ipu.org/wmn e/world.htm](http://www.ipu.org/wmn%20e/world.htm)
- Jacobs, M., Kent, A., & Watkins, K. (2003). *Progressive global ization: Toward an international social democracy*. London: Fabian Society.
- Judt, T. (2009). What is living and what is dead in social democracy. *New York Review*, 59(20), 86–96.
- Keynes, J. M. (2007). *The general theory of employment, interest and money*. Basingstoke, UK: Palgrave Macmillan. (Original work published 1936)
- Kitschelt, H. (1994). *The transformation of European social democracy*. Cambridge, UK: Cambridge University Press.
- Krugman, P. (2010, January 10). Learning from Europe. *The New York Times*. Retrieved June 1, 2010, from <http://www.nytimes.com/2010/01/11/opinion/11krugman.html?scp=1&sq=Learning%20from%20Europe&st=cse>
- Lenin. (1901). *What is to be done*. Retrieved May 28, 2010, from <http://www.marxists.org/archive/lenin/works/1901/witbd/i.htm>
- Marquand, D. (1993). After socialism [Special issue]. *Political Studies*, 41.
- Marshall, T. H. (1950). *Citizenship and social class and other essays*. Cambridge, UK: Cambridge University Press.
- Marx, K., & Engels, F. (1848). *The communist manifesto*. Retrieved May 28, 2010, from [www.marxists.org/archive/marx/works/1848/communist manifesto/index.htm](http://www.marxists.org/archive/marx/works/1848/communist%20manifesto/index.htm)
- Organisation for Economic Co operation and Development. (2010). *OECD Factbook*. Washington, DC: OECD.
- Paterson, W. E., & Thomas, A. H. (Eds.). (1986). *The future of social democracy*. New York: Oxford University Press.
- Polanyi, K. (1962). *The great transformation*. Boston: Beacon Press (Original work published 1944)
- Przeworski, A. (2001). Socialism and social democracy. In J. Krieger & M. E. Crahan (Eds.), *The Oxford companion to politics of the world* (2nd ed., pp. 775–780). Oxford, UK: Oxford University Press.
- Sasson, D. (1995). *One hundred years of socialism*. New York: New Press.
- Steger, M. (Ed.). (1996). *Selected writings of Eduard Bernstein, 1900–1921*. Atlantic Highlands, NJ: Humanities Press International.
- Steger, M. (1997). *The quest for revolutionary socialism*. Cambridge, UK: Cambridge University Press.
- Steyn, M. (2009). Prime minister Obama. *National Review*, 61(5), 22–24.
- Tilton, T. (1990). *The political theory of Swedish social democracy*. Oxford, UK: Clarendon Press.
- Tudor, H., & Tudor, J. M. (Eds.). (1988). *Marxism and social democracy*. New York: Cambridge University Press.

78

LENINISM, COMMUNISM, STALINISM, AND MAOISM

ANNA N. MAKHOVA-GREGG

Austin Peay State University

Communist ideology, in the form of its various “brands” (Leninism, Stalinism, and Maoism), has had a powerful impact on shaping political realities throughout the 20th century. In fact, the most consequential political events of the past century can neither be explained nor understood without a clear reference to communist ideas and the most significant attempts at their implementation. It is important to understand that the political slate was not just wiped clean with the turn of the millennium. The need for furthering scientific analysis of the communist ideology and the variety of its implementations certainly warrants including this chapter in the *21st Century Political Science* handbook.

Today, the term *communism* is most often used with reference to either the theory by Karl Marx and Friedrich Engels or the politico-economic regimes that claimed to use Marxian theory as their foundation. This chapter has four objectives. The first is to briefly summarize the most essential principles and concepts of the original theory as developed by Marx and Engels. The second objective is to outline two interpretations of the Marx–Engels theory, one by Lenin and one by Stalin. The third objective is to compare and contrast the teachings of Mao Zedong¹ with the Leninist ideology. Fourth, some of the theoretical lessons from the Soviet and Chinese experiences with communism will be discussed, along with their possible implications for the international political landscape of the 21st century.

The Marx–Engels Theory of Communism

The original ideas about eliminating social inequalities and creating a perfectly egalitarian society can be traced back many centuries to ancient Greece (e.g., Hesiod, 1985; Plato, 2006) and to medieval Europe (e.g., Campanella, 2007; More, 2002). Some of these fragmented ideas were finally assembled in the form of a relatively coherent theory by two 19th-century German thinkers, Marx and Engels. Their theory was developed as an intellectual reaction to the socially painful side effects of the 19th-century industrial revolution in advanced Western economies. Their approach encompassed philosophic, political, and economic components that were borrowed from numerous social theories abandoned in preindustrial Europe. Among the thinkers who are known to have had the most powerful influence on Marx and Engels, and who therefore should also be given credit for their contribution to classical Marxism, are German philosophers such as Immanuel Kant, Georg Hegel, and Ludwig Feuerbach and British political economists such as Adam Smith and David Ricardo, as well as a large group of French social theorists including Jean-Jacques Rousseau, Charles Fourier, and Henri Saint-Simon, along with many others.

The term *classical Marxism* is widely used as a reference to theories, concepts, and ideas expressed in the original works by Marx and Engels. This set of ideas is a very broad and very complex theoretical framework and not a

simple single coherent theory, as has sometimes been portrayed by self-proclaimed followers. Also, it is obvious that the generalizations extracted by Marx and Engels from their analysis of 19th-century capitalism in Europe could not easily be used to explain economic and political realities of other times and places. These generalized and dated philosophies have invited a multitude of various interpretations and adaptations, three of which are discussed further in this chapter. Marx and Engels were very prolific writers, and their works take up several volumes of very dense and scrupulous technical prose that would be impossible to properly summarize in this chapter. Therefore, the few theories and concepts outlined here were selected on the basis of their importance for the three ideological brands previously noted.

The theoretical teachings of Marx and Engels contain ideas that can broadly be divided into two general categories. The first is devoted to the critique of the existing socioeconomic regime and attempts to justify the inevitability of capitalism's demise. The ideas in the second category are directed toward developing a futuristic model of a distinctly modern, fully egalitarian, and therefore more just politico-socioeconomic order. This side of the classical Marxist teachings is rather diluted and full of ambiguities and even contains occasional contradictions. This regime of the future was labeled *communism* by the authors and stems from the term that originated in revolutionary France in the 1840s. Communism, according to Marx and Engels, would supplant capitalism through the series of social revolutions initiated in the industrialized West and eventually spread throughout the world. Ironically, instead of mobilizing industrial workers within economically advanced countries, the ideas of Marx and Engels ignited revolutionary movements among two preindustrial agrarian societies in the East, first in Russia and later in China.

Two manuscripts stand out among the most influential works by Marx and Engels: the *Manifesto of the Communist Party* (1848/2002) and *Capital, Volume I* (Marx, 1867/1992).² The former piece, which is often referred to as *The Communist Manifesto*, is essentially a small brochure of approximately 40 pages that outlined the program of the Communist League, an international organization established in Paris in 1836 as the *League of the Just* and formally disbanded in 1852. *The Communist Manifesto*, however, acquired a life of its own and is still referred to as one of the most influential political manuscripts written. In this work, Marx focuses his analysis on the problems of capitalism and lays out the concept of the *class struggle*. There are few details in *The Manifesto* with regard to the specific form that communism would take as it replaced the brutal, unfair, and fundamentally controversial capitalism.

The other manuscript critical to understanding the theory of Marx and Engels is *Capital: A Critique of Political Economy* (Marx, 1992, 1993). This document covers some 1,400 pages and is full of specific economic terms and

formulas. The first volume of *Capital* focuses on analysis of capitalist economy, its origins, future, and structural contradictions and the resulting class struggle between workers and owners. Throughout this volume, Marx only vaguely hints at the specifics of how the new *just economic order* that is supposed to replace capitalism will operate once the industrial working class (the *proletariat*) frees itself from being exploited by the owners of the means of production (the *bourgeoisie*). The second and third volumes, which were intended to detail the theory of the post-capitalist method of production, are essentially a collection of Marx's drafts, sketches, and fragments, edited and compiled by Engels after Marx's passing. The fact that *classical Marxism* offered a compelling critique of capitalism but lacked a coherent theory of *scientific socialism* had attracted numerous interpretations and adaptations not unlike Leninism, Stalinism, and Maoism.

The idea of inevitable transition from capitalism to socialism historically stems from an evolutionary process driven by economic development and is the approach that Engels labeled *historic materialism*. Having been influenced by Darwin's theory of evolution, Marx and Engels viewed the development of human society as a progressive succession of politico-economic regimes (*economic formations*), wherein each succeeding regime is superior to the preceding one in terms of *production relations*. The concept of production relations pertains to ownership, distribution, and redistribution of product and constitutes the *base*. Borrowing the materialistic approach of Ludwig Feuerbach and the dialectical approach of Georg Hegel, the founders of Marxism believed that the base is the driving force that eventually leads to change in the *superstructure*, a particular form of social consciousness that includes legal, political, and cultural institutions that reflect the base. This idea that throughout history in every society, the base has always determined superstructure is often referred to as *economic determinism*. Together, Marx argued, the base and the corresponding superstructure determine the *mode of production*, which, in turn, defines the *economic formation*, a developmental stage in the history of humankind.

Historic materialism is a theoretical perspective of social, political, and economic development that views the history of humankind through the lens of economic determinism. This approach to history portrays the perpetual class struggle between workers and owners over ownership of the *means of production* as the main driving force behind societal progress. The term *means of production* refers to physical, nonhuman inputs used in production, such as factories, machines, and tools. According to Marx and Engels, the concept of class struggle plays a central role in explaining society's alleged inevitable development from economic oppression under capitalism to a classless and propertyless society in which the means of production are owned by an entire society of equals.

Formulation of the theory of social evolution inspired by Charles Darwin's *Origins of Species* and analysis of

contemporary capitalism were only a prelude to what Marx and Engels claimed was their main discovery: the finality of the class society. According to their teachings, capitalism is the last of its kind, that is, the last in the succession of economic formations based on economic exploitation. The collapse of capitalism as a result of a proletarian revolution will become the beginning of a new, highest, and final stage in the development of humankind, the creation of a perfectly just society in which labor is reconciled with the means of production and, therefore, economic exploitation is no longer possible. According to Marx, it would also mean the end of history.

To conclude this brief overview of classical Marxism, it might be useful to mention the main features that made this theoretical framework so attractive to its followers. One of the main appeals of the theory of Marx and Engels to their followers was its ability to explain (and capitalize on) the painful social dislocations indicative of the industrializing capitalist societies of 19th-century Europe. Some of the predictions with regard to capitalist economies appeared to be coming true, such as Marx's claim that the contradictory nature of the capitalist economy leads to periodic crises. According to Marx, over time these crises would become more protracted and eventually would become fatal to capitalism itself. Although it is true that industrialized countries seem to be prone to periodic crises, none of them has (as of yet) resulted in a social breakdown.

The forward-looking character of Marxian predictions with regard to the future establishment of the perfectly egalitarian, classless, and propertyless social order has certainly encouraged many people in countries such as czarist Russia and postimperial China who struggled with industrialization in part because of their traditional backward-looking culture. Also, the international or *ethnoneutral* character of communist teachings, which imply that the main irreconcilable contradictions are between proletariat and bourgeoisie, not between ethnic groups, lends itself as a solid foundation for unifying people of various origins under one leadership.

Communist Ideology in Russia and the Soviet Union: Leninism and Stalinism

Leninism has been a dominant branch of Marxism for most of the 20th century. This offshoot of communist ideology was named after Vladimir Lenin,³ who was the mastermind of the Russian revolution in 1917 and became the founder of the Soviet Union in 1922. Unlike theory-inspired classical Marxism, Leninism developed as a result of practical efforts to apply the teachings of Marx and Engels to Russian conditions. Lenin's manuscripts, therefore, are focused on practical solutions to specific problems of organizing a successful revolution and building a socialistic economy and state.

One of the most significant departures of Leninism from the doctrines of Marx and Engels was its claim of

revolutionary potentials for Russia's peasantry, the poor who represented the rural population and were largely employed in traditional agricultural production. To justify the inevitability of socialist revolution in Russia at the beginning of the 20th century, where industrial labor still constituted a small minority of the working class, Lenin had to get creative. In his well-known manuscript titled *What Is to Be Done?* (1902/2002), published 15 years prior to the Revolution of 1917, Lenin proclaimed that peasantry is essentially the *agrarian proletariat*, which could be inspired by professional revolutionaries (such as Lenin himself) to join the industrial workers in deposing the imperial government, taking power, and establishing the *dictatorship of the proletariat*,⁴ the temporary state that facilitates transition from capitalism to communism.

To describe Russia's peasants at the end of the 19th century as the agrarian proletariat was not just conceptual stretching but an outright subversion of the truth. The majority of them were loyal to the monarchy and the Orthodox Church and, as of 1916, owned 89.1% of the agricultural land in European Russia. This does not mean, however, that peasants were satisfied with the regime. Traditional agricultural production under the difficult climate conditions, paired with a growing rural population, was failing to sustain the livelihood of peasant communities, causing chronic food shortages and periodic famines during the years when the climate was especially unfavorable. Episodes of civil unrest were severely suppressed by the extreme autocratic regime, which maintained a tight grip over the vast Russian Empire through its extensive bureaucratic apparatus, the police, and the army.

An interesting and not widely known, but well-documented, fact about Lenin's passion for his revolutionary activities was that he was motivated, not by sympathy for the poor, but rather by his hatred for the existing social and political order in Russia (Pipes, 2001). Born into the family of a high-ranking civil servant who was awarded hereditary nobility, Lenin had personal reasons to become embittered toward the regime, which in 1887 executed his older brother, Alexander, for involvement in a conspiracy to assassinate Czar Alexander III Romanov and expelled young Lenin from the University of Kazan for participation in a minor student disturbance, ruining his hopes for a career as a lawyer.

Another significant departure from the teachings of Marx and Engels was Lenin's original idea that a *vanguard party* formed of professional revolutionaries, who did not belong to the working class themselves and whose revolutionary aspirations, therefore, would be untainted by the trade union mind-set, would lead the proletarian revolution. This departure had consequences that reached far beyond the Russian and Chinese revolutions to inspire followers of Marxist ideology later in the 20th century, such as Ernesto "Che" Guevara and Fidel Castro.

The idea of the vanguard party as an architect of the social revolution became closely related to the central

doctrine of Leninism: *bolshevism*.⁵ This doctrine represented Leninism's most significant departure from the teachings of Marx and Engels because it proposed to organize the party in a military-like fashion with a strict chain of command, a membership requirement of full-time commitment to revolutionary activities, and a top-down approach to the working masses, who needed to be "educated" and guided toward a violent power takeover. Bolshevism, as it was developed by Lenin in the years prior to 1917, essentially represented a very selective approach to classical Marxism. Unlike Marx, who put the main emphasis on *natural* historical development driven by economic determinism, Lenin argued that Russia did not have to undergo "bourgeois" revolution and should not allow capitalism to develop fully before it would be ready for a socialist revolution and the dictatorship of the proletariat.

In the early 20th century, many European supporters of the ideas of Marx and Engels were wondering why capitalism had not collapsed in accordance with the main prediction of the socialist dogma. Lenin's answer to this question came in the form of a new theory. In 1916 and 1917, he wrote a book titled *Imperialism: The Highest Stage of Capitalism* (1999), in which he argued that advanced capitalist countries, in their futile attempts to avoid perpetual economic crises, engage in colonizing or imposing economic dependence on less developed countries in Asia, Africa, and Latin America. By gaining access to the markets in these countries and exploiting their labor and natural resources, the imperial powers try to prop up their ailing economies and buy off their industrial workers. Lenin insisted that this strategy was yet more evidence of capitalism's decay and simply postponed the inevitable change to socialism.

Russia's economic situation worsened dramatically in 1914 as it became involved in World War I. Rampant inflation and food and fuel shortages, paired with rumors of government inefficiency and corruption, created a great deal of internal discontent on the part of peasants, industrial workers, and intellectuals from the middle and upper classes. The autocratic rule of Czar Nicholas II was rapidly losing its legitimacy among all layers of the Russian society, and Lenin's position on war was uncompromising. Lenin claimed that war between the nations had to be turned into a war between the social classes. Workers should turn their weapons against their exploiters and place their power in the hands of the only party that claimed to truly represent their interests, the Bolsheviks.

Observing the devastating effects that World War I had on Russia's economic and social life, Lenin developed yet another reality-inspired theory, which he called the *theory of the revolutionary situation*, in which he further discussed one of his fundamental works, *The State and the Revolution* (Lenin, 1917/1993). According to the theory of the revolutionary situation, three conditions must be present for a revolution to be successful: (1) a profound crisis within the powers that be, (2) unusual hardships suffered by the working

people, and (3) a sharp spike in social unrest and political involvement by the masses. Despite the fact that Russia, during the years 1916 and 1917, met all three of these conditions, the revolution that happened in early March of 1917 was nothing that Lenin had expected or wanted. Hunted by the Russian imperial police, he watched from abroad as the events in Russia unfolded and did not return to Petrograd⁶ until July of 1917.

After the Russian army, which was poorly supplied and staffed primarily with a group of rebellious peasant draftees, suffered a series of defeats on the German front, Czar Nicholas II was pressured by his generals to abdicate his power in order to save Russia from defeat. Political power was then assumed by a group of parliamentary deputies, who called themselves the *Provisional Government*. The socialist-minded opposition within the Russian parliament, together with politically active Petrograd intellectuals, created a concurrent institution, the *Soviet*. This was a council of workers and soldier representatives who intended to serve as a "watchdog" over the actions of the Provisional Government. These developments created a regime of dual power under which the Soviet relentlessly criticized and undermined the authority of the Provisional Government without being held responsible for the consequences of its decisions and actions. The collapse of the autocratic monarchy, which had imposed unity on the Russian Empire for centuries through a combination of traditional legitimacy and forceful oppression, submerged Russia into anarchy.

In the fall of 1917, the Provisional Government lost all support from the Russian army's leadership, and the Bolsheviks won a majority in the Petrograd Soviet. Lenin worked hard to convince his followers and other factions in the Soviet to seize the opportunity and take power, and as a result, a coup took place on November 7, 1917.⁷ To disguise the seizure of power by one party and himself as the leader of this party, Lenin put forth the slogan "All Power to the Soviets," which promised the relinquishing of state authority to the chain of newly established grassroots organizations throughout Russia (the Soviets), which at the time attracted the loyalties of the working masses and rebellious soldiers. The first two pieces of legislation written by Lenin and adopted by the new regime were the *Decree on Peace* and the *Decree on Land*. The former announced Russia's immediate withdrawal from World War I and its concession to most of Germany's territorial demands. The latter abolished private ownership of land and announced a redistribution to the peasantry of land owned by nobility, the Orthodox Church, and the monasteries. As naive as that legislation was, it, along with other freedoms (press, religion, the formation of political organizations, etc.), evoked widespread support for the Bolsheviks among the uneducated masses of peasants, the war-tired soldiers, and the small layers of industrial workers who, at the time, represented less than 2% of Russia's population.

There is minor disagreement among scholars with regard to the motivation behind the next several steps undertaken by the Bolshevik-led government. Some of the scholars, whose views are well articulated by Richard Pipes (2001), argue that the dictatorship and totalitarian regime established in Russia in 1918 represented the original intent of Lenin and his supporters. Another group of researchers and Lenin's biographers, however, maintain that Lenin's ideal of a workers' democratic state was shattered against the harsh reality of civil war and foreign intervention that engulfed Russia shortly after the Bolshevik-controlled Soviets took power (Hesli, 2007). Whatever Lenin's original intent was, in 1918 in his famous work *The Proletarian Revolution and Renegade Kautsky* (Lenin, 1969), he completely discarded any possibility of achieving the communist ideals through a peaceful democratic process, conveniently forgetting that in their last years, Marx and especially Engels came to recognize that possibility.

For more than 2 years after the November revolution, Russia was engulfed by civil war. Coupled with the Bolsheviks' War Communism—a program of forced nationalization, grain requisitions, and labor mobilization—the civil war of 1918 to 1921 resulted in economic devastation and consolidation of political power in the hands of the Soviets. Succumbing to the increasing threat of massive peasant uprising and the Kronstadt Revolt, the 10th Congress of the Communist Party, guided by Lenin, inaugurated the New Economic Policy (NEP). This was the beginning of a period when the extremely coercive Bolshevik state was transformed into a much less violent set of regulatory institutions that finally were able to establish a relative monopoly on the use of force. During this brief interlude between the civil war of 1918 and Stalin's Great Purge of the late 1930s, Russia began acquiring some basic features of a modernizing state.

The NEP was essentially a package of economic policies that provided for basic economic freedom of enterprise in strictly limited areas and aimed at giving a chance for the economy to recover, resume growth, and end famine. According to Lenin himself, the NEP represented a tactical temporary retreat toward capitalism. Essentially the newly established Soviet state was forced to make a political compromise in order to survive the devastation brought on by civil war and the politics of War Communism. The Soviets loosened restrictions on small private industrial enterprises and substituted requisitions of agricultural produce from the peasants for a tax with a rate known in advance. It is important to point out that in spite of significant economic freedoms during the NEP years, the state maintained full political control and kept exclusive hold on the commanding heights of the economy, including finance, large and medium industry, modern transportation, foreign trade, and wholesale commerce. These policies stimulated the economic recovery in agriculture and small manufacturing, as well as the state-controlled industries. The temporary drift toward partial decentralization

did not just allow the Russian economy to recover but also provided an opportunity for the Communist Party—now the sole power holder—to regroup and shift gears from a struggle for power during the civil war to peacetime governance.

The first Soviet constitution of 1918, drafted by Lenin, created the Russian Soviet Federated Socialist Republic, which territorially was just a shadow of the vast Romanov empire. A quasi-federalist structure of the Soviet Union, which included Russia, Ukraine, White Russia (now Belarus), and Transcaucasia (Armenia, Azerbaijan, and Georgia), was established on December 30, 1922, and sealed by the second Soviet constitution in 1924. This arrangement was the result of a Lenin-inspired compromise between the Bolshevik desires for strong central control and the national independence movements in the borderlands. This new arrangement resurrected czarist policies of *Russification* toward ethnic minorities such as Tatars, Bashkirs, Adygs, Cherkas, Chukchas, and many others (more than a hundred in total; McAuley, 1992).

Stalin consolidated power after Lenin's death in 1924 and quickly proceeded to build the basis for his own autocracy by crushing his real and even potential political opponents inside the party. He abruptly discontinued any economic freedoms granted by the NEP and moved to full centralization of the state, accompanied by massive repression of various population groups. Essentially, the public sector of the economy was consuming the private sector, first through industrialization of manufacturing and then through the massive, forceful collectivization of agriculture. The private sector quickly disappeared as the majority of the Soviet citizens had become state employees by being forced to join collective and state farms or being sent to the numerous forced labor camps (Pipes, 2001).

Massive centralization of the Soviet state took place during the last decade before World War II. Many regulatory state agencies were quickly changing into repressive machinery, which grew significantly in size as the number of citizens classified as *enemies of the people* reached millions. The most prominent among the governmental agencies involved in purges was the political police, known since 1934 as the Narodny Komissariat Vnutrennih Del (better known by its initials, NKVD, and which translates into English as People's Commissariat of Internal Affairs), which guarded and administered forced labor camps. The culture of fear formed during the Great Purge of the 1930s created a foundation for what appeared to be a total control of the state (with Stalin as the head of state) over society. This control was achieved through mobilization of the country in a quasi-military manner and at the cost of great suffering by the Soviet people. Although some population groups were affected by particularly acute purges, prosecutions affected virtually all Party organizations, government branches, and the army (Conquest, 1985).

One of the most prominent aspects of Stalin's legacy became the phenomenon subsequently labeled his *cult of personality*. Essentially, it was a subculture that portrayed

Stalin as an omnipresent, omnipotent, and infallible god-like figure and that remained intact through his death in 1953. One of the many explanations of the cult-of-personality phenomenon common to most Communist regimes has to do with the fact that in premodern traditional societies (such as Russia and China in the first half of the 20th century), it was natural to attribute divine qualities to political leaders. This property has been frequently exploited by the Communist leaders themselves to make up for the lack of legitimacy of their totalitarian states.

Stalin's brand of Communism acquired yet another peculiar feature, which set it apart even further from classical Marxism and to some extent from Leninism as well. Stalin was first among Communists to attempt to capitalize on nationalist sentiments and xenophobia to inspire passionate compliance among the masses and to promote fearful obedience among minorities. Classical Marxism viewed *nationalism* as one of the tricks that the bourgeoisie used to deflect the proletariat from forming a unified front in its quest to put an end to social injustices and economic exploitation. Lenin early in his political career saw nationalism as a hindrance to the destruction of the old regime, but later he attempted to embrace it as one of the state-building tools. Stalin, however, recognized nationalism's superior potential to unify masses of certain ethnic descent. Compared with vague ideas of international Communism, nationalism and especially xenophobia became a much more efficient way of appealing to the raging emotions of the majority and inspiring fear and slavish compliance among the minorities. As time went on, Stalin aligned himself with Russian chauvinism,⁸ aiming his repressive governmental machine at various minorities in Russia, especially the Jews. Very quickly he realized that this political posture also offered an outlet for blaming any failure of his despotic regime on ethnically distinct groups within the population. Many of the Communist leaders since Stalin have attempted and often succeeded in playing the nationalist card to exploit its centrifugal potentials to justify their oppressive regimes.

Stalin's reign before and after World War II became the time when the Soviet state acquired the general structure and long-standing traditions that remained essentially the same until its demise in 1991 (Kryshtanovskaya, 2004). The 1936 Stalin constitution was not revised until 1977, despite significant changes in the economy, state-society relations, internal political climate, and international affairs. Even four decades after Stalin's constitution was written, the new edition of the constitution written under the leadership of Leonid Brezhnev still did not contain any provisions requiring a major reorganizing effort. Rather it emphasized continuing "glorious" traditions of the past. Comparative literature in Soviet politics points toward a number of factors contributing to this continuity of the Soviet regime. Many of these factors have to do with various aspects of widespread corruption in the Soviet political system, which had its origins in the first few decades after the Union of Soviet Socialist Republics was established

and which reached its peak under Brezhnev (Kaminski, 1989). The establishment of a single-party system during Lenin's period shielded the Communist Party from the possibility of facing any external political opposition and effectively disabled electoral institutions. Stalin's purges did away with the party's internal opposition, any effective division of powers, and plurality of opinions within the society. Brezhnev's "contribution" to this political culture of institutional decay and disregard for the written laws was a system of patronage based on personal loyalties and unlimited tenure in office at all levels of party and state hierarchies. In essence, if Stalin resurrected the czarist autocracy, Brezhnev reintroduced the decision-making methods of premodern elements into the state organizations (Simis, 1982).

Chinese Experience With Communism: Mao Zedong's Marxist-Leninist Orthodoxy

According to Isaac Deutcher (1966), a well-known and respected British historian, the communist revolution in China of 1947 can be classified as one of the *ironies of his tory*. "Lacking any native ancestry [with regard to socialist or communist philosophy], Chinese Communism descends straight from Bolshevism. Mao stands on Lenin's shoulders" (Deutcher, 1966, p. 90). This statement reflects a generally accepted notion among Western researchers, as well as among the followers of *Chairman Mao*, that Maoism is a direct descendant of Marxism-Leninism. However, a group of scientists who looked beyond the general similarities argued that Maoism departs from Leninism as Leninism departs from Marxism (Meisner, 1971, 1982; Pfeffer, 1976). Essentially, what makes Maoism an offshoot of Leninism and not a direct descendant of Marxism is its pragmatism or willingness to modify and even subvert theory in the face of realities and the inclination to accept that the ends would always justify the means. According to the researchers that view Leninism and Maoism as deviants from Marxism, Lenin and Mao took Marx's ideas out of a time- and space-relevant political context and attempted to forcefully fast-forward history into the distant future that Marx had predicted.

Born in 1893 in the family of a wealthy Chinese farmer, Mao Zedong had always displayed an emotional attachment to and naive fascination with the traditional Chinese peasantry. In fact, both aforementioned groups of Chinese political researchers share an understanding that the question of the peasant's role in the communist revolution is the main deviation of Maoism from its Marxist-Leninist origins. This issue is crucial to an understanding of the Chinese brand of communism because it has a direct impact on the central concepts of this ideology, such as *the class struggle* and the relationship that should exist between the working class and its leading party. Lenin saw the Communist Party as the main source of revolutionary consciousness destined to save the proletariat from what he called the *trade union*

mind set. He defined this concept of trade union mind-set as the willingness to change political and economic institutions to be more favorable to the working class through evolutionary rather than revolutionary means. Maoism puts even more emphasis on revolution as the only instrument suitable to achieve the true change of a regime. The most striking evidence of this was Mao's policy of *continuous revolution*, which he put forth 10 years after coming to power in 1947.

Mao and his followers also adopted three other key elements of Lenin's version of Marxism. The first was the Marxian explanation of historical development (economic determinism) as amended by Lenin with the concept of the Communist Party as the sole agent of change toward the final stage of politico-economic development. Chinese communists also borrowed much of Lenin's practical recommendations with regard to the vanguard party's organizational principles, such as *democratic centralism*.

The second crucial feature of Leninism integrated into the core set of Maoist writs is Lenin's theory of imperialism as the highest and final stage of capitalist development. With the breakup of the colonial system after the end of World War II, assisting anti-imperialist forces in the third world appeared to be a significantly more feasible foreign policy to both communist countries—the Soviet Union and China—at the time. As their relationship with each other deteriorated, their resolve to sway the former colonies in Asia and Africa toward a specific brand of communism (the Soviet or the Chinese) became almost as intense as their desire to steer them away from the capitalist path of development. North Korea (and later Vietnam) became the primary battlefields for the ideological struggle. Mao's ambitions in the third world brought very modest rewards at best. The People's Republic of China could not afford to devote the necessary (and ever-increasing) quantities of economic resources required to pursue his costly and far-reaching ambitions. Third world leaders, who were willing to auction off their countries' developmental paths to the highest bidder, lacked devotion to a specific ideology and were easily overthrown by domestic political forces.

The third important element that links Maoism and Marxism–Leninism has to do with the conviction that to prevent counterrevolution, the communist regime had to engage in ruthless demolition of the entire institutional structure of the previous regime. However, Mao Zedong took this rather extreme principle even further by initiating assault on the newly established institutional structure of the People's Republic of China during a crusade labeled the *Great Proletarian Cultural Revolution* (1966–1968). This crusade targeted any existing form of authority (government, party, family), with the sole exception of Mao's personal leadership, and essentially was undertaken in retaliation for the failed attempt to overcome economic backwardness during the so-called *Great Leap Forward* (1958–1961), which claimed millions of lives and submerged

China's economy into chaos. This succession of socio-economic experiments did not come to a halt until Mao's death in 1976.

Communist Ideology in 21st-Century Politics

The three most consequential brands of communist ideology surveyed here have posed many questions and presented many issues to historians and political researchers. In an attempt to draw political lessons from the Soviet and Chinese experiences, it is important to remember that, like any political phenomenon, communism emerged in response to time-specific political processes such as the industrial revolution during the second half of the 19th century. It had the explicit intent of communicating the political interests of industrial workers, whose basic needs were neglected by the existing politico-economic order, to the power-holding political elite.

When discussing the implications of communist ideology for the political landscape of the 21st century, one cannot avoid making analytical distinctions between the original *idea* of creating a perfectly egalitarian, economically just society, a political *program* put forth by Marx and Engels, and the numerous attempts at practical *implementation*. Most likely, the communist idea will continue to emerge in a variety of ideological forms as long as there are economic disparities, which can be perceived as injustices by various individuals, social groups, and even whole nations. The Marxist political program, which was created specifically to address social dislocations of 19th-century industrialization in advanced Western countries, had certainly become outdated in the 20th century, which prompted a number of political entrepreneurs, including Lenin, Stalin, and Mao Zedong, to reinterpret and rebrand it. Finally, the discussed earlier attempts to establish regimes based on the communist idea and the program of Marx and Engels have had a lasting effect on the power distribution within the international system. The emergence and later disappearance of the Soviet Union as one of the two global superpowers resulted in a unipolar world, which according to some experts is a much less stable arrangement than the bipolar world of the cold war era. Today's China is often characterized as a rising global economic power although it is still debated whether this economic wonder came about because of communist rule or in spite of it.

The question that many experts on communist and post-communist politics have been trying to answer is whether the failure of communism was a result of faulty attempts at its implementation or an unrealistic societal model that was doomed to fail from the time it was first conceived. Even though opinions are mixed with regard to assigning blame, many scholars favor the latter view. To quote Pipes (2001), "Communism was not a good idea that went wrong;

it was a bad idea” (p. 147). But some scholars are quick to point out that each and every one of the communist regimes was forced, in the face of political realities, to depart significantly from both the utopian communist ideal and the dogma of Marx and Engels in order to survive as long as it did. Having said that, however, the interests that brought this idea to life and made it popular enough to have a lasting and irreversible impact on millions and millions of lives cannot be simply dismissed from the political stage. Even though virtually all the communist regimes failed before the end of the same century in which they had emerged, there is no guarantee that the quest for perfect social equality is over.

The implosion of the Soviet regime became a textbook example of the ideological bankruptcy of Marxism–Leninism. Stalinism had discredited itself and decomposed even prior to the Soviet collapse. Even though today’s People’s Republic of China still claims the title of a communist regime, its practice of *socialist market economy* is largely based on private ownership of the means of production and has much more in common with leading capitalist economies than with other past or existing communist regimes.

Conclusion

The four major goals outlined at the beginning of this chapter were to (1) summarize Marxian theory, (2) position the ideological interpretations of Lenin and Stalin with respect to the original Marxian teachings, (3) compare the Maoist philosophy with the major premises of Marxism–Leninism, and (4) evaluate the sociopolitical implications of the experiments with communist ideals in Russia and China for the political landscape of the 21st century. The first section of the chapter focused on the major premises of the theoretical works of Marx and Engels. It highlighted the sketchy character of the Marxian vision of the future, which leaves ample room for further interpretations and study by anyone willing to test it empirically by conducting social experiments. The second section of this chapter examined the two versions of such social experiments conducted by Vladimir Lenin and subsequently by Joseph Stalin in the Soviet Union during the first half of the 20th century. The focus of the third section was the Chinese interpretation of communist ideals by Mao Zedong in the second half of the 20th century. A special effort has been made to distinguish between the original Marxian ideas, the teachings of Lenin, and the “insights” of Mao. Finally, the fourth section aimed at assessing the ideological impact of communist ideology and its practical applications on the political landscape of the 21st century.

Since the November Revolution of 1917 in Russia, there have been dozens of attempts throughout the world to establish regimes based on various interpretations of the

communist principles formulated by Marx and Engels. Virtually all of them have failed. The three most consequential attempts for world politics were those of Leninism and Stalinism in Russia and Maoism in China.

Most researchers conclude that Marxism, the theoretical foundation of communism, contained the seeds of its own demise. Based on faulty views of history, unrealistic economic foundations, and unscientific psychological doctrines, this ideology had a power to inspire but lacked coherence to deliver on its promises. Every one of the numerous attempts to use it as a blueprint had to rely on ruthless coercion, imposing enormous social, economic, and psychological costs on the societies that bought into a Marxist vision of the future.

It is important to understand that none of the social self-proclaimed followers of the communist ideology diligently followed all the guidelines outlined by Marx and Engels. Even the most basic premise that transition to communism would take place in economically advanced industrialized countries was violated in every one of the attempts to apply the Marxian doctrine. Also, Marx and Engels could not possibly be specific enough and could not predict the enormous number of specific issues that their practice-driven disciples were encountering. These ambiguities within Marxian dogma spawned a number of ideological “brands” that in various degrees departed from the original, but all of them had to rely heavily on coercion.

To sum up the features of the aforementioned ideologies that qualified them as interpretations of Marxism one would have to mention the contention of private property and the notion of human nature as fundamentally malleable through coercion and education, as well as superficial emphasis on the material side of societal existence. Of the three attempts to build communism, Leninism was the most willing to accommodate realities of political, economic, and social processes, as the discussion of the NEP has illustrated. Stalinism was slightly more practice driven, possibly because of Stalin’s limited talents for theoretical thinking. The cult of personality, xenophobia, and morbid bureaucratization, typical to a certain extent of all the communist regimes, were taken by Stalin to an extreme level. Divine status of the *beloved leader* and encouraged disregard for any other form of authority became the most prominent features of Maoism.

Notes

1. Because of transliteration, different spellings of this name have been proposed (e.g., Mao Tse Tung, and Máo Zédōng). Mao Zedong is commonly used in the most recent literature.

2. *Capital, Volume II*, and *Capital, Volume III*, were posthumously published by Engels in 1885 and 1894, respectively.

3. Vladimir Ilyich Ulyanov (1870–1924) is better known under his revolutionary pseudonym, Lenin, which he was using to escape prosecution by the Russian imperial police for his revolutionary activities prior to 1917.

4. This term was first used by Marx in *Critique of the Gotha Program* (1875).

5. *Bolshevik* can be translated literally from Russian as “majority.” The term was coined by Lenin’s supporters, who constituted the majority after the split of the Social Democratic Party in 1903. The party members who disagreed with Lenin had to accept the title *Mensheviks*, meaning “minority.” They argued that socialist revolution had to originate from a grassroots movement within the working class.

6. In August 1914, as Russia entered into war with Germany, the name of the Russian capital was changed from the Germanic *St. Petersburg* into the more Russian equivalent, *Petrograd*. After the death of Lenin, the city was renamed once again as *Leningrad* in 1925. The original name of *St. Petersburg* was restored in 1991.

7. Prior to 1918, Russia followed the Julian calendar, which was 13 days behind the Gregorian calendar, which the rest of Europe used. The Bolsheviks’ revolution took place on the night of November 7 (Gregorian calendar), which is equivalent to October 25 (Julian calendar), and this explains why the November coup is sometimes called the “October Revolution.”

8. This was ironic because Stalin himself was ethnically Georgian and not specifically Russian.

References and Further Readings

- Campanella, T. (2007). *City of the sun*. New York: Cosimo.
- Conquest, R. (1985). *Inside Stalin’s secret police: NKVD politics, 1936–1939*. Stanford, CA: Hoover Institution Press.
- Deutcher, I. (1966). *Ironies of history*. London: Oxford University Press.
- Fitzpatrick, S. (1999). *Everyday Stalinism: Ordinary life in extra ordinary times: Soviet Russia in the 1930s*. New York: Oxford University Press.
- Hesiod. (1985). *Works and days* (T. Sinclair, Ed.). London: Georg Olms Verlag.
- Hesli, V. (2007). *Government and politics in Russia and the post Soviet region*. Boston: Houghton Mifflin.
- Kaminski, A. (1989). Coercion, corruption, and reform: State and society in the soviet type socialist regime. *Journal of Theoretical Politics*, 1, 77–102.
- Kryshtanovskaya, O. (2004). *Anatomy of the Russian elite*. Moscow: Solov’ev.
- Lenin, V. (1969). *Lenin reader: The outstanding works of V. I. Lenin*. Washington, DC: H. Regnery.
- Lenin, V. (1993). *The state and the revolution*. New York: Penguin. (Original work published 1917)
- Lenin, V. (1999). *Imperialism: The highest stage of capitalism*. Sydney, Australia: Resistance Books.
- Lenin, V. (2002). *What is to be done?* Cambridge, UK: Cambridge University Press. (Original work published 1902)
- Luxemburg, R., & Wolfe, B. D. (1981). *The Russian revolution and Leninism or Marxism?* Westport, CT: Greenwood Press.
- Mao Tse tung. (1990). *Quotations from chairman Mao Tse tung*. San Francisco: China Books.
- Marx, K. (1992, 1993). *Capital: A critique of political economy* (Vols. 1, 2, 3). New York: Penguin. (Volume 1 originally published 1867)
- Marx, K., & Engels, F. (2002). *The communist manifesto*. New York: Penguin.
- McAuley, M. (1992). *Soviet politics 1917–1991*. Oxford, UK: Oxford University Press.
- Meisner, M. (1971). Leninism and Maoism: Some populist perspectives on Marxism Leninism in China. *China Quarterly*, 45, 2–36.
- Meisner, M. (1982). *Marxism, Maoism, and utopianism: Eight essays*. Madison: University of Wisconsin Press.
- More, T. (2002). *Utopia*. New York: Cambridge University Press.
- Pfeffer, R. (1976). *Mao and Marx in the Marxist Leninist tradition*. *Modern China*, 2(4), 421–460.
- Pipes, R. (2001). *Communism: A history*. New York: Modern Library.
- Plato. (2006). *The laws*. Sioux Falls, SD: NuVision.
- Safire, W. (2008). *Safire’s political dictionary*. Oxford, UK: Oxford University Press.
- Simis, K. (1982). *U.S.S.R.: The corrupt society: The secret world of Soviet capitalism*. New York: Simon & Schuster.
- White, S. (2001). *Communism and its collapse*. New York: Routledge.

79

SOCIALISM IN THE DEVELOPING WORLD

LILLY KELAN LU

University of North Texas

This chapter reviews the literature on the adoption and development of socialism in developing countries. First, it presents the definition and basic features of socialism. Second, it reviews theories that attempt to explain why underdeveloped countries adopt socialism, discussing specific countries as examples. Third, it addresses the challenges faced by the third world in transitioning from socialism. Fourth, it reviews the variants of classical Marxism in the third world, with emphasis on the differences between socialism in developing countries and classical Marxism. Fifth, the chapter discusses the potential benefits and problems of revised socialism in the developing world.

Why Did the Developing World Adopt Socialism?

Socialism, according to Baradat (1997), is defined as a system that grants the ownership of production to the public and provides the public with a social welfare system while pursuing material abundance, equality, and sharing for its people. The *third world* refers to the countries that were “peripheral to the center of world capitalism and subordinated to [it] through colonialism or various forms of imperialist or ‘neo-colonial’ control and penetration, and where indigenous capitalism was weakly developed” (White, Murray, & White, 1983, p. 4).

Various scholars (Baradat, 1997; Desfosses & Levesque, 1975; Elliott, 1962; Fagen, Deere, & Coraggio, 1986; Kautsky, 1968; Kurzman, 1963; White et al., 1983) have investigated the reasons socialism was adopted by the developing world. Generally, the reasons can be divided into four categories. First, the nature of socialism emphasized immediate economic growth, which could help the third world develop within a relatively short period. Second, the nature of socialism was anti-imperialism and antiexploitation. This reflected the third world’s radical response to hostile threat and even military aggression from the global imperialists and its economic dependence on the developed world. The third reason is a political-cultural argument that claims that the traditional societies in the third world were more compatible with socialism. Fourth, the Soviet Union’s efforts to influence developing countries led to the development of socialism in such countries.

Economic Development

First, the adoption of socialism reflected the desire of the intellectuals in the third world to explore a model for improving economic development. According to Kautsky (1968), it was the modern elite (or intellectuals) in the third world, and not the general population, that was attracted to communist models. In contrast, traditional elites (including aristocratic rulers, big landowners, and

the clergy) tended to regard communist models as a threat, and the masses (including peasantry, urban petite bourgeoisie, and a modern working class) were not broad-minded enough to understand the communist models. The modern elites, as a group, absorbed their values, not from their native traditional society, but from other industrially advanced societies. The adopted values mainly include an emphasis on enhancing the economy; increasing wealth; and improving social equality, education, and political participation. Further dispersion of these values started prevailing in developing countries via universities, armies, bureaucracies, and trade unions.

Second, although the goals set by the intellectuals in the third world were adopted from the Western world, according to Desfosses and Levesque (1975), the Western way of achieving these goals was slow compared with the communist model. The modern elite believed the socialist model was more applicable because it provided an economic growth plan that promoted much faster economic development.

Third, communist models promoted rapid industrialization under the lead of the intellectuals (Kautsky, 1968). Communist regimes share similarities with the regimes in other developing countries because in both cases, intellectuals led similar social sections, had similar opponents, and were pursuing similar values. Another reason the socialist model appealed to intellectuals in the third world is the potential opportunities it offered for the intellectuals to obtain power and prestige.

Anti-Imperialism and Anti-Exploitation

Another reason for the developing world to adopt socialism is the nature of socialism, anti-imperialism, and anti-exploitation. First, the third world countries faced the problem of economic dependence on developed countries (Desfosses & Levesque, 1975; Fagen et al., 1986). Most of the smaller markets of the third world relied on exports of cash crops and/or natural resources and were thus more vulnerable to changes in international markets. The people in the developing world had strong nationalistic desires for independence and self-determination and for economic and political development. Socialism satisfies the need of developing countries to reverse their lack of development and to strive for freedom.

Second, socialism served as a model for third world people and leaders to deal with their ambivalence toward industrialization (White et al., 1983). In these countries, socialism not only provided hope of gaining the benefits that were achieved by capitalism and industrialization, but it also eliminated the exploitation that was prevalent in the capitalist model. Moreover, socialism assured that the people's representatives represented the people's best interests.

The Compatibility Between Traditional Societies and Socialism

Another reason some developing countries adopted socialism was that these countries perceived socialism as being compatible with traditional societies (Desfosses & Levesque, 1975). This compatibility mainly focused on the nonexistence of classes. According to classic Marxism, the essence of socialism is that the victory of the proletariat is caused by class struggle. Many intellectuals in the third world claimed that there were no issues of class in their countries after independence. Moreover, these intellectuals argued that a proletariat was not a necessity for realizing socialism. They also argued that in the developing world, the essential issue is to mobilize the general public for political activities instead of depending on class struggle led by the proletariat.

Soviet Influence

The adoption of socialism in the third world is due not only to the appeal of socialism itself but also to Soviet efforts to influence developing countries in order to win support against the Western world. Kurzman (1963) provided some evidence for this argument. First, Communists' emphasis on the third world did not start until the end of World War II. After the war, the Communists' influence in the underdeveloped countries ramped up. The major achievement by the Communists was the adoption of communism in China in 1948. Communism became established in North Vietnam, Cuba, Laos, South Vietnam, Indonesia, and the Indian state of Kerala. In the Middle East, communism was set up in Syria in 1955 and Iraq in 1959. In Africa, communism was adopted in Guinea, Ghana, the formerly Belgian-controlled Congo, and Kenya.

What Were the Barriers to the Socialist Transition in the Developing World?

According to Fagen et al. (1986), socialist transition movements pursue three major goals. The first goal is to produce and redistribute sufficient wealth to satisfy the people's basic needs. The second goal is to provide equal opportunity to the people in pursuing improvement in income, justice, and culture. The third goal is to reestablish the relationship between the state and the whole society and thus mobilize the public for political participation.

For Fagen et al. (1986), there are three barriers to socialist transition in the third world: the characteristics of the international system, the particular difficulties faced by countries with small and peripheral economies, and the scarcity of suitable models for the third world's transitions.

The first problem that influences the third world's transition is the global system (Fagen et al., 1986). To begin

with, the global system was dominated by capitalist states during the end of the 20th century. Capitalists wrote the rules of global games such as trade, finance, and investment. For the socialist countries, it was almost impossible to participate in these global games while isolating themselves from global prices, markets, technology, and tastes that were established by the capitalist countries. Although *multipolarity* (i.e., the presence of more than two roughly equal superpowers who dominate world politics) played a secondary role in providing the third world with opportunities such as political, military, and ideological leverage, multipolarity was not the fundamental arena of third world development. For example, with respect to issues such as trade, finance, technology, investment, and foreign aid, the world was more of a unipolar system than a multipolar one. Moreover, the world's unipolar feature was reflected in the cultural domain. In the realm of consumption, demand by the countries in the world became increasingly homogeneous (e.g., demand for Coca-Cola and blue jeans), which reflected the sole dominance of advanced capitalism.

The second problem is the particular difficulties faced by countries with small and peripheral economies (Fagen et al., 1986). To begin with, the development of these small and peripheral economies was firmly constrained by the unipolar system. These economies had no capacity to set prices in trade, to invent technology, or to accumulate their own domestic savings. Therefore, they had to depend heavily on the developed countries. Moreover, the small and peripheral economies have been influenced deeply by the culture of developed countries. The third world's traditional culture has been penetrated by Western culture and the Western lifestyle. The changing nature of culture made it difficult for the third world to bridge the gap between its own traditional culture and socialism.

The developing countries' efforts to get access to the international core market and to strive for a better export environment have been offset by increasingly more strict business barriers in the developed world for the protection of their local economy. Due to low domestic savings, third world countries needed low-cost and loosely restricted loans. However, the advanced, industrialized countries preferred and bargained to lend to developing countries in order to maximize their own profits. In sum, the highly developed countries occupy much more advantageous positions in the global market than do small and peripheral countries.

Given the harsh international political and economic environment for smaller and peripheral economies, it was even more difficult for those smaller and peripheral economies that decided to transition to socialist regimes. These socialist-transition countries bore two heavy burdens. First, similar to nontransition countries, socialist-transition states needed to struggle for survival and/or success in economic development and cultural preservation. Second, they needed to make fundamental transformations

in their domestic political, economic, and social systems in order to pursue social justice. Moreover, the countries undergoing fundamental transformations in their socialist transition were faced with political, economic, and military pressures from the advanced, industrialized, capitalist countries, who regarded the socialist transition in smaller and peripheral economies as a threat to their established rules. The threat perceived by the capitalist countries was understandable because socialist transitions usually happened in groups and tended to ally with established socialist countries. More important, these socialist transitions challenged the founding logic of capitalism and thus threatened the dominant position of capitalism in the world. Besides the barriers set by the imperialists and their allies within the socialist countries, socialist transition movements faced potentially intrinsic problems (Fagen et al., 1986). First, although the old global system was under attack, it was an integrated and stable economic and political system, which was important for wealth accumulation. However, socialist transition movements interrupted that stability while being accompanied by class struggle, which was not beneficial for economic growth unless a strong dictator could keep the chaos in control.

The third barrier was a lack of practical socialist theory for the third world countries to refer to during their transitions (Fagen et al., 1986). Although Marxist theory analyzed the conditions that would assist socialist transitions, Marxism did not explicitly illustrate how to transition to socialism and what socialism was. Because different countries had different situations and experiences, it was difficult to apply the models promoted by Mao or Lenin to other developing countries.

The Variants of Socialism in the Developing World

According to Baradat (1997), since Marx's death, three major varieties of socialist movement developed. The first variant is the *orthodox school*, which defended Marx's work against any significant revisions. The second variant is the *revisionists* and the *Fabians*, who attacked the major Marxist theories and preferred a smooth and peaceful approach to achieving the goals of socialism. This school has been adopted mostly by European countries and the United States. The third variant is Marxism–Leninism. Based on Marxism, the ideology proposed by Lenin explained the reasons for wrong predictions in Marxism. Moreover, Marxism–Leninism emphasized the guidance of an elite group because of the lack of class consciousness among the proletariat. The elite group would play a leading role in eliminating rebellion after the fall of the capitalist system. Because the proletariat would replace the bourgeois rulers and continue to grow, the proletariat would eventually become the sole economic class in the society, which would approach the ideal plan in classic Marxism.

The reasons third world leaders chose their respective variants of socialism in their adoption of elements of Marxism, Leninism, or Maoism has been analyzed by Desfosses and Levesque (1975) and Fagen et al. (1986) as follows. The first reason was the third world intellectuals' nationalistic determination to avoid subjugating their country to absolute control of socialists after getting rid of the long-term absolute control by the imperialists (Desfosses & Levesque, 1975). Second, the continuing failure of capitalism in smaller and peripheral economies was another reason socialist regimes were attractive to the third world (Fagen et al., 1986). The failure of capitalism to satisfy the material needs and provide social justice for people in the third world made socialism an attractive alternative. Moreover, increasingly people in the third world realized that the world was under the control of the advanced capitalist countries, which set global rules and reaped as many benefits as possible from the third world. In contrast, socialism promised to eliminate the old global rules and to bring social justice and rationality through policies such as public ownership and centralized economic planning. Third, leaders in the third world thought the socialist model would speed up their economic growth and allow their countries to eventually catch up with the industrialized capitalist countries (Desfosses & Levesque, 1975). The goal of pursuing economic development in the third world had superseded the goal of realizing socialist ideology. Fourth, people in the third world desired to preserve their traditional culture and national characteristics and strove to minimize the interruptions and distortions of industrialization and modernization (Desfosses & Levesque, 1975).

Because of these reasons, intellectuals in third world countries were selective when adopting classical Marxism, Leninism, or Maoism. Therefore, no uniform socialist model was observed among all the developing countries but rather a variety of versions of socialism.

The major differences between socialism in the developing world and classic Marxism have been investigated from two major perspectives. White et al. (1983) focused on the differences between revolutionary socialism in the third world and the classic Marxism. Kautsky (1968) investigated the differences between the Soviet model and the goals of intellectuals in underdeveloped countries in the post-World War II period.

The differences between revolutionary socialism in the third world and classic Marxism have been summarized by White et al. (1983) as follows. First, revolutionary socialism succeeded in the underdeveloped countries but not in advanced capitalist countries as predicted by Marx. Second, socialism in the third world was not the heir of capitalism but rather was a historical alternative. Third, socialism in the third world was not unifying and consolidating the international working class among the well-developed industrial countries but rather was facilitating radical nationalism in underdeveloped countries. Fourth,

revolutionary socialism was not established on the cultural and economic basis of highly developed capitalism but rather drove accelerated, delayed development under unfavorable conditions at home and abroad. Fifth, the power of revolutionary socialism did not depend on the proletariat but rather on existing classes and strata, mainly including the peasantry and the petite bourgeoisie.

More specifically, Kautsky (1968) and Baradat (1997) investigated the differences between the Soviet model and the models followed by the developing countries. First, within the classic model, the proletarian revolution of the communists' Marxian symbolism did not appeal to the developing countries. In spite of some changes in communist symbols in the Soviet model after World War II, "revolution" and "proletarian revolution" were still regarded by the third world as important communist symbols. However, some policies in the third world were opposed to revolution and proletarian revolution.

Second, although at the beginning the classic model appealed to intellectuals in the third world, inflexible obedience to Marxian doctrine and related policies created in developing countries an image of communists as proletarian revolutionaries. Moreover, this image made them appear to be attempting to win proletarian supporters and searching for capitalist enemies in the Western world. This actually weakened the appeal of the Soviet model in the developing countries. However, because of its incompatibility with actual developments in the third world, Marxian doctrine and related policies started to decay in the third world while the appeal of the Soviet model increased.

Third, although the classic Soviet model appealed to the intellectuals who led anticolonial movements in developing countries, these intellectuals were very different from Soviet intellectuals. The intellectuals in the third world were a strong prorevolutionary group. Although they were committed to building a strong and new society, they were afraid of deep industrialization because of their fear of losing power. The successful industrialization in the Soviet Union led to the replacement of revolutionary intellectuals with managerial intellectuals.

Case Studies for Six Socialist Developing Countries

Six case studies further illustrate the reasons the third world adopted variants of socialism and the characteristics of these variants. The cases are Cuba, Iraq, China, Yugoslavia, Libya, and North Korea.

Cuba

Several scholars (Baradat, 1997; Bideleux, 1985; Desfosses & Levesque, 1975; White et al., 1983) have used Cuba as a case study to investigate its variant of socialism. Socialism was adopted by Cuba under Fidel

Castro and his followers to build socialism with Cuban national characteristics. By taking into consideration Cuba's relevant historical and international experience, Castro and his followers foresaw a break with the United States after a radical Cuban revolution. In spite of the close interdependence between Cuba and the United States, Castro and his followers regarded the intertwining of the United States and Cuba as a barrier to any substantial structural transformation of Cuban society. After the failed U.S. invasion of the Bay of Pigs in April 1961, Castro and his followers transformed, step by step, their domestic and international programs simultaneously, achieving a complete transition to socialism and a turning toward the Soviet Union.

Moreover, the political leaders perceived that some of the principles of Marxism–Leninism fit quite well with the values gained from their struggle against ousted Cuban president Fulgencio Batista. These values included *activism*, *voluntarism*, and *elitism*. Voluntarism in particular has been emphasized in socialist ideology with Cuban characteristics.

In addition, Cuba's Communist Party played an important role in allying the country with the Soviet Union. The alliance with the Soviet Union equipped Cuba with important political strategies. In spite of this, the essence of Cuban socialism was not provided by the Soviet Union.

Before adopting socialism, Cuba was a country whose public was not only accustomed to extensive government intervention but also very active in political participation. The introduction of socialism brought about several changes. First, government expanded its intervention by controlling the ownership of all the means of production. Second, to a certain extent, the new government both encouraged more political participation and tried to keep the participation under the control of the government.

Cuba's socialism had its own characteristics. First, the development of a revolutionary Communist Party of Cuba was late and slow. According to Baradat (1997), Castro did not become a Marxist–Leninist until he came into office. Communism as an ideology and Cuba's Communist Party as an organization were used by Castro as an instrument to control the country. Therefore, the Communist Party itself in Cuba did not enjoy as much authority as that enjoyed in communist countries such as Soviet Union and China. In contrast, Cuban socialism was built on the personal charisma of Castro, which became a barrier to the future transfer of leadership through the Communist Party system. Second, the Cuban government encouraged the widespread use of economic collectivization, with emphasis on agricultural production, especially sugar production. Third, Cuba used moral and nonmaterial incentives to motivate the public employed in production. Nevertheless, these moral incentives were closely related to promises of health care and retirement benefits and priority in accessing political positions.

Iraq

Socialism began developing in Iraq in 1948 when Iraq had increased contacts with socialist countries. The deepening of socialism in Iraq was fueled by nationalism (Desfosses & Levesque, 1975). With increasingly negative feelings in the Arab world toward the West, more and more Iraqis regarded the West as exploitive and preferred socialism over capitalism in Iraq. In addition, some of the ideas and values of socialism were incorporated into Iraqi nationalist ideology, including nationalist issues with respect to Palestine, the Israeli issue, and the unity of the Arab world.

The leading party among the Arab nationalist parties in promoting socialism was the Baath party. The Baath's notion of socialism evolved from utopian communism to a mix of Marxist scientific socialism and nationalism. The Baathist approach to establishing socialism in Iraq had its own characteristics and thus was called *Arab socialism*, which indicated that it was not derived from Marxism but rather from a contrasting ideology. First, instead of using the term *revolution* in their struggle to build socialism, Baathists used the term *coup* and regarded it as the only way to achieve the renaissance of nationalism and the establishment of socialism in Arab countries. Second, although the Iraqi constitution claimed ownership of all property belonging to the Iraqi people, equal distribution of all economic resources among the people, and the state's control of the means of production, private property rights were not outlawed but were instead protected. All Iraqi citizens enjoyed the freedom to own real estate as long as the amount of property owned by a citizen did not exceed the amount used directly by the citizen and was not used in a way that exploited others. Third, under the leadership of the Baathists, social class differences were to be eliminated and a more fair and equal social order was to be established.

China

The socialist variant in China—Maoism—was admired by a large number of developing countries (Baradat, 1997). Maoism was an attempt to adjust classic Marxism–Leninism to Chinese traditional culture, agrarian economy, and guerrilla war (Baradat, 1997). According to White et al. (1983), Maoism reflected China's break with Eastern European models of state socialism, which represented either hierarchical bureaucratism (e.g., the Soviet Union) or market socialism (e.g., Yugoslavia). In contrast, Maoism was built on the initial socialist values, including equality, participation, and collectivism. The Chinese model under the guidance of Maoism was promoted by development experts as a good example for other developing countries to follow. According to Baradat (1997), compared with classic Marxism–Leninism, Maoism has several unique characteristics.

The first distinguishing feature of Maoism is *populism*. Mao and his followers emphasized the detrimental role of peasants in the victory of the Chinese revolution. In contrast, classic Marxism identified the proletariat as the leading force of communist revolution. In order to reconcile Maoism with classic Marxism, Mao highlighted the virtues of the peasants, such as purity, simplicity, and flawlessness, as the fundamental strengths of the Chinese people. Moreover, during the Cultural Revolution, Mao forced urban sophisticates to work in rural farms for the purpose of “learning from the people” (Sobhe, 1982, p. 273).

Second, Maoism emphasized *ideological purity* over economic training. This characteristic of Maoism revised the economic determinism of classic Marxism. According to Mao, peasants can be proletarianized through both proletarian mentality education and economic experience. However, this transformation of the peasants did not require them to leave their land.

The third feature of Maoism is *permanent revolution*. In spite of the vague references to the definition of permanent revolution by Marx and Lenin, Mao, more radically, emphasized permanent revolution instead of sustaining a status quo. According to Mao, permanent revolution and constant violence were needed along the road to socialism because great progress was born from social disorder. In Maoist logic, socialism and capitalism can never coexist peacefully.

The fourth feature of Maoism is its *rejection of elitism*, specifically Lenin’s dependence on the Communist Party to lead the revolution. This principle arose from a concern that the fruits of the Chinese revolution might fall into the hands of bureaucracies instead of the people. Therefore, Mao emphasized that the people were the ones who should be trusted to attain the final goals of the revolution under the guidance of communist ideology because the people were red intrinsically. By thus mobilizing the people, China underwent several campaigns, such as the antilandlord campaign (1949–1952), the first 5-Year Plan (1953–1957), the Hundred Flowers Campaign (1957), the Great Leap Forward (1958–1960), and the Great Cultural Revolution (1966–1976).

The fifth principle of Maoism is its *tolerance of the bourgeoisie*. The major reason for this tolerance was the poor economic situation in China during the early years of communist power in 1948 and 1949. Instead of sticking to the classic Marxist doctrine of socializing the economy, Mao learned the lesson from Lenin that immediate socialization of the economy is dangerous. Rather, the bourgeoisie could contribute to the stabilization of the Chinese economy. The real enemies were those who exploited the Chinese people, mainly landlords and imperialist capitalists. Within Maoism, anti-imperialism was closely incorporated into Chinese nationalism.

The last distinguishing feature of Maoism is guerrilla warfare. Classic Marxism and Leninism stressed the short duration of revolutions. In contrast to Marx’s anticipation of spontaneous revolution and Lenin’s conspiratorial

revolution (his use of secret operations, among other tactics; Lee, 2003), Mao expected an extended period of revolutions in the developing world.

Yugoslavia

Yugoslavia’s variant of socialism—Titoism—was born of Yugoslav leader (Josip Broz) Tito’s resistance against the Soviet Union’s strong pressure to force Yugoslavia to conform (Baradat, 1997). In spite of his original firm support for Stalin, Tito eventually led Yugoslavia to break out of the Soviet bloc in 1948 and benefit from trade with both the Western and the Eastern bloc.

Although Tito split from the Soviet bloc, he continued to be committed to Marxism (Baradat, 1997). However, he made some revisions to classic Marxism–Leninism. In contrast to Lenin’s stress on violence and Marx’s and Lenin’s idea of permanent revolution, Tito’s version of communism de-emphasized the use of violence in socialist development. According to Tito, as soon as socialism developed to an advanced phase, violence was not needed. This made Tito much more moderate than the classic communists. Furthermore, Tito promoted coexistence and active cooperation with capitalist countries.

According to Tito, different countries have different interpretations of socialism (Baradat, 1997). Tito emphasized that the threats to socialism came not from imperialism but from socialist countries’ overly centralized domestic systems. In contrast to Marx, Tito realized that overcentralization tended to cause the growth of bureaucracy and the possible exploitation of the masses.

In order to resolve the problem of overcentralization, Tito decentralized Yugoslavia’s political and economic organization, a design called *market socialism*. First, he encouraged private ownership and control, as well as the decentralization of industry. According to Tito, although workers would directly control the factories in the most advanced stage of socialism, the nationalization of industries was an important first step in reaching that stage. In the future, advanced socialist societies should be based on social control rather than state control. Second, Tito decided to decentralize the dominant rule of the Communist Party. According to Marx, political institutions would eventually disappear in a classless society, leaving only an economic bureaucracy. Tito did not agree with Marx on this point but insisted that the Communist Party should always be the leader of the system in spite of the absence of the need for centralized economic control. Third, Tito also decentralized the governmental structure by dividing the country into six republics and two autonomous provinces, pushing the government toward federalism.

Libya

Socialist revolution in Libya was brought about by Mu‘ammar Muhammad al-Gadhafi and his followers in

1969 after overthrowing the king of Libya (Desfosses & Levesque, 1975). According to Gadhafi, the socialist revolution in Libya was guided by two basic values: social justice and Islamic socialism. The adoption of socialism in Libya, as elsewhere, involved revisions to classic Marxism–Leninism. The major basis for Libya’s revisions was Gadhafi’s belief in the compatibility between socialism and the fundamental Islamic philosophy. He thus regarded socialism as an indigenous political ideology in Libya rather than one adopted from the outside world. Moreover, the approach to private property rights and widespread nationalization valued in classic Marxism–Leninism were perceived by Gadhafi as in accordance with Muslim tradition because Islamic law endorsed private property rights and nationalization without damage or injury. Despite these compatibilities between the Libyan socialist model and the classic communist model, Libya’s model also exhibited some important variations.

First, the development of socialism in Libya was based on two basic values, social justice and modernization. According to Desfosses and Levesque (1975), socialism in Libya was regarded more as a value system than as an economic system because the socialism pursued in Libya aimed to realize the true dignity of human beings through political participation and through ensuring an acceptable standard of living. Libyan socialism regarded socialism as the fruit of class cooperation, rather than class struggle as suggested by classic Marxism–Leninism.

Second, Libya’s history and economic situation contributed to Gadhafi’s special economic policies (Desfosses & Levesque, 1975). Because of the country’s low level of economic development, domestic capitalists in Libya were not considered as an enemy to be eliminated. In addition, because of Libya’s experience of exploitation by foreign imperialist powers, Gadhafi differentiated between domestic and foreign capitalists. Foreign capitalists, because of their exploitative nature, were regarded as incompatible with Libya’s national interests and were nationalized by the government. Domestic capitalists, on the other hand, were not regarded as enemies of the people unless they were exploitative. Thus, under Libyan socialism, citizens were allowed to own capital privately and to invest their capital freely as long as they were not exploitative.

Third, Gadhafi realized that the small Libyan bureaucracy lacked the education, expertise, and manpower necessary to modernize the nation (Desfosses & Levesque, 1975). As a result, nationalization is not the whole answer to either imperialism or lack of development. Gadhafi envisioned a Libyan agricultural and industrial economy that would achieve the complete potential of Libya’s economy while ensuring the control of the people over the basic means of production. Libya’s petroleum wealth should be used to modernize the country and improve the life of all its citizens, not to enrich foreign capitalists or a few Libyans.

North Korea

North Korean socialism was brought about by Kim Il Sung, who was the leader of both the state and its Communist Party until his death in 1994 (Baradat, 1997). North Korea’s pattern of socialism was not the same as that of the Eastern European countries that followed the Soviet Union. The special characteristics of North Korean socialism were mainly reflected in the extreme ideological, political, and economic isolation of North Korea.

Because Marxism–Leninism was accused of not being able to solve all the problems in Korea, Kim Il Sung did not allow anyone except himself to read the classic works by Marx, Lenin, and Stalin. Moreover, Kim Il Sung prevented the public from accessing *any* outside information (Baradat, 1997). Through language reform beginning in 1945, North Koreans were cut off from their history because younger generations could not read texts written before 1945. Because they were not informed about North Korea’s liberation from the Japanese by the Soviet Union or the assistance provided by the Chinese army in the war against the United States in 1950, people in North Korea firmly believed the ideology that Kim Il Sung was the sole leader of world revolution.

Economically, North Korea carried out a harsh self-dependence policy, called *juche*, defined as “a line of economic construction for meeting by home production the needs for manufactured goods and farm produce necessary for making the country rich and strong and improving the people’s labour and one’s own national resources” (White et al., 1983, p. 129). Under this policy, North Korea isolated itself almost entirely from the world economy. In spite of its severe lack of two important resources, oil and coal, North Korea pushed the public to work overtime and developed technology transformations to structure its national economy so as to minimize its dependence on and vulnerability to the external world for the long term.

Third, North Korean socialism was built within a militarized society (Baradat, 1997). Since the Korean War, from 1950 to 1953, North Koreans had maintained a permanent militarization against South Korean and U.S. armies. The armed forces in North Korean society were used as an instrument of socialization. Meanwhile, family and educational institutions were also used as important instruments of socialization. Therefore, the entire North Korean society became militarily oriented.

Problems of the Variants of Socialism in the Developing World

The first problem in some variants of socialism lies in the combination of nationalism and socialism. According to White et al. (1983), although nationalism helped the third world catch up with the development of industrialized countries, intense nationalism sometimes turned into *chauvinism* (as in Pol Pot’s Cambodia) and *exclusionism* (as in North

Korea). Moreover, White et al. (1983) found that the context of military threat and conflicts before and after revolution led to nationwide militarization both ideologically and institutionally and to an intense level of security consciousness. This problem accounted for the stagnation of “socialist internationalism” and the frequent conflicts among socialist countries.

Second, several scholars have found that it was problematic for some developing countries to establish socialism on the foundation of their traditions. Desfosses and Levesque (1975) suggest that the issue of combining peasants’ collective traditions and socialist agricultural collectivization not only contradicted some classic doctrines of Marxism, Leninism, and Maoism but also slowed agricultural development in third world countries. According to Desfosses and Levesque (1975), the similarity between the peasants’ communal tradition and socialist agricultural collectivization had been exaggerated by third world socialists. For example, in Africa, the communal tradition among peasants was mainly reflected in their cooperation in consumption but not in production. However, it was cooperation in production that was needed to make communal agriculture the foundation of development. Therefore, building socialism in Africa on the peasants’ communal tradition actually violated some basic principles of classic Marxism, Leninism, and Maoism. Moreover, even if socialists in Africa tried to change cooperation in consumption into cooperation in production, there was still no guarantee that the agricultural sector would provide sufficient surplus for the development of industry. Rather, radical changes in the system of production, such as changing the traditional systems of land ownership, kinship, and duties within a community, actually slowed down agricultural development.

Third, although the goals (such as equality, social justice, and the nationalization of industries) pursued by third world socialists helped third world countries accumulate social support, the disappointment among the public for any failure to reach those goals posed a grave threat to socialist regimes (Fagen et al., 1986). Moreover, because the theories of socialism emphasized its advancement in industrialization and higher efficiency, the public in third world socialist countries tended to overestimate the industrial capability of their countries and to de-emphasize the importance of the peasants.

Fourth, according to Kurzman (1963), among underdeveloped countries, communists share similar strategies: *left strategy* and *right strategy*. Under the left strategy, communists in each country would fight against noncommunists and accuse them of assisting imperialism. In contrast, the right strategy was usually applied when there was a common threat to both communists and noncommunists. The communists would ally with noncommunists to fight the common enemy, as the Soviet Union had done when it was threatened by Nazi Germany before the end of World War II. In most underdeveloped countries, communism was spreading by

first using the right strategy to cooperate with noncommunists in defeating common enemies and then using the left strategy to eliminate the noncommunists within the country. This is the tactic of deception described in Kurzman (1963). Using this approach, communist propaganda tended to subordinate the Marxist doctrine to the fight against “Western imperialism.”

Future Directions

According to Baradat (1997), the fall of Marxism–Leninism, Stalinism, Titoism, and Maoism does not mean that socialism itself will fall. Some scholars have claimed that there is recent evidence of the possible return of socialism in the former Soviet Union and in eastern Europe (Baradat, 1997). Baradat suggests, first, that after experiencing free education, universal medical care, and social protection, people from many socialist countries tend not to give them up, given that not many countries are able to deal with the uncertainties and insecurities of individualism. Second, because government can play an important role that cannot be filled by individuals, socialism has a future with its goal of a decent standard of living for every person, along with protection by the state from predation by economically stronger members of society.

References and Further Readings

- Abdel Malek, A. (1968). *Egypt: Military society*. New York: Vintage.
- Baradat, L. P. (1997). *Political ideologies: Their origins and impact*. Upper Saddle River, NJ: Prentice Hall.
- Bideleux, R. (1985). *Communism and development*. New York: Cambridge University Press.
- Desfosses, H., & Levesque, J. (Eds.). (1975). *Socialism in the third world*. New York: Praeger.
- Draper, T. (1965). *Castroism: Theory and practice*. New York: Praeger.
- Elliott, J. R. (1962). *The appeal of communism in the underdeveloped nations*. Dubuque, IA: W. C. Brown.
- Fagen, R. R., Deere, C. D., & Coraggio, J. L. (Eds.). (1986). *Transition and development: Problems of third world socialism*. New York: Monthly Review Press.
- Griffith, S. (1961). *Mao Tse tung: On guerrilla warfare*. New York: Praeger.
- Haim, S. (1964). *Arab nationalism: An anthology*. Berkeley: University of California Press.
- Hinton, H. C. (1973). *An introduction to Chinese politics*. New York: Praeger.
- Jacobs, M., Kent, A., & Watkins, K. (2003). *Progressive globalization: Toward an international social democracy*. London: Fabian Society.
- Judt, T. (2009). What is living and what is dead in social democracy. *The New York Review of Books*, 59(2), 86–96.
- Kaple, D. A. (1994). *Dream of a red factory: The legacy of high Stalinism in China*. New York: Oxford University Press.

- Kautsky, J. H. (1968). *Communism and the politics of development: Persistent myths and changing behavior*. New York: Wiley.
- Khadduri, M. (1970). *Political trends in the Arab world: The role of ideas and ideals in politics*. Baltimore: Johns Hopkins University Press.
- Kitschelt, H. (1994). *The transformation of European social democracy*. Cambridge, UK: Cambridge University Press.
- Krugman, P. (2010, January 10). Learning from Europe. *The New York Times*. Retrieved May 26, 2010, from <http://www.nytimes.com/2010/01/11/opinion/11krugman.html?scp=1&sq=Learning%20from%20Europe&st=cse>
- Kurzman, D. (1963). *Subversion of the innocents: Patterns of communist penetration in Africa, the Middle East, and Asia*. New York: Random House.
- Lee, S. (2003). *Lenin and revolutionary Russia*. Hong Kong: Routledge.
- Marquand, D. (1993). After socialism [Special issue]. *Political Studies*, 41.
- Marshall, T. H. (1950). *Citizenship and social class and other essays*. Cambridge, UK: Cambridge University Press.
- Marx, K., & Engels, F. (1848). *Manifesto of the Communist Party*. Retrieved May 26, 2010, from www.Marxists.org/archive/marx/works/1848/communist-manifesto/index.htm
- Moody, P. R. (1983). *Chinese politics after Mao*. New York: Oxford University Press.
- Paterson, W. E., & Thomas, A. H. (Eds.). (1986). *The future of social democracy*. New York: Oxford University Press.
- Post, K., & Wright, P. (1989). *Socialism and underdevelopment*. London: Routledge.
- Przeworski, A. (2001). Socialism and social democracy. In J. Krieger & M. E. Crahan (Eds.), *The Oxford companion to politics of the world* (2nd ed., pp. 775–780). Oxford, UK: Oxford University Press.
- Rauch, G. von. (1972). *A history of Soviet Russia* (6th ed.). New York: Praeger.
- Robinson, C. J. (1983). *Black Marxism*. Chapel Hill: University of North Carolina Press.
- Rozman, G. (Ed.). (1992). *Dismantling communism*. Washington, DC: Woodrow Wilson Center Press.
- Sasson, D. (1995). *One hundred years of socialism*. New York: New Press.
- Short, M. I. (1993). *Love and religion in Marxist Cuba*. New Brunswick, NJ: Transaction.
- Singleton, F. (1976). *Twentieth century Yugoslavia*. New York: Columbia University Press.
- Sobhe, K. (1982). Education in revolution: Is Iran duplicating the Chinese Cultural Revolution? *Comparative Education*, 18(3), 271–280.
- Steger, M. (Ed.). (1996). *Selected writings of Eduard Bernstein, 1900–1921*. Atlantic Highlands, NJ: Humanities Press.
- Steger, M. B. (1997). *The quest for evolutionary socialism*. Cambridge, UK: Cambridge University Press.
- Suarez, A. (1967). *Cuba: Castroism and communism*. Cambridge: MIT Press.
- Tilton, T. (1990). *The political theory of Swedish social democracy*. Oxford, UK: Clarendon Press.
- Townsend, J. R. (1974). *Politics in China*. Boston: Little, Brown.
- White, G., Murray, R., & White, C. (Eds.). (1983). *Revolutionary socialist development in the third world*. Lexington: University Press of Kentucky.
- Woodward, S. L. (1995). *Socialist unemployment: The political economy of Yugoslavia*. Ewing, NJ: Princeton University Press.
- Yan, S. (1995). *The Chinese reassessment of socialism*. Ewing, NJ: Princeton University Press.

PART VI

AMERICAN POLITICS

80

FOUNDING OF THE AMERICAN POLITICAL SYSTEM

RICHARD J. HARDY

Western Illinois University

The American political system did not suddenly appear. It has evolved through a unique and complex convergence of countless civic actors, political ideas, and significant events. To grasp the complexities of our unique political system, it is important to understand the roots of our polity, how it has emerged, and some of the major theories concerning its development. This chapter begins with discussion of the major elements that influenced the founders of our republic. Those influences include the early colonial experience, the British legal system, natural rights philosophy, classical republicanism, and religious teachings. Next, the chapter examines the sequence of events and underlying factors leading to the American Revolution, the Declaration of Independence, and the Articles of Confederation. Attention finally turns to the United States Constitution and its framers, underlying concepts, theories, and remarkable features.

Major Influences on American Constitutional Development

Influence of British Legal History

The American political system was obviously influenced by British legal history. The government of Great Britain is not based on a single written constitution. Rather, the British legal system is composed of an amalgam of

political documents, common law, customs, and acts of the Parliament. The most influential documents in the development of British government are the Magna Carta (1215), the Petition of Rights (1628), and the English Bill of Rights (1689). The following is a brief description of each of these landmark documents and how they influenced subsequent American political documents.

The Magna Carta (1215), or Great Charter, is the foundation on which modern British government is built. Originally, England was ruled by an absolute monarch, one who had virtually unrestrained control over his or her subjects. In 1215, the feudal barons forced King John to concede a list of rights to the English nobility. The Magna Carta's 63 chapters are provisions calling for limited government and the preservation of certain personal and property rights. Chapter 39, for example, stipulates, "No free man shall be taken or imprisoned . . . except by the lawful judgment of his equals or [and] by the law of the land." The clause "judgment of his peers" refers to trial by jury, the right later guaranteed by the U.S. Constitution's Sixth Amendment, and the idiom *law of the land* becomes the foundation for the concept of due process of law, a phrase later found in the U.S. Constitution's Fifth and Fourteen Amendments.

Monarchical concessions were further made in 1628, when the Parliament forced King Charles I to sign the Petition of Rights. This document contains a list of grievances that parallels many of the complaints chronicled in

the Declaration of Independence and an assortment of rights contained in the U.S. Constitution (1787). For example, the Petition of Rights declares that the monarch may not (a) arrest legislators when attending Parliament (Article I, Section 6.1 of the U.S. Constitution provides a similar immunity for members of Congress); (b) force civilians to house and feed soldiers (the Third Amendment of the U.S. Constitution contains a similar proscription); (c) levy taxes without approval from the House of Commons (Article I, Section 7.1 likewise mandates that revenue bills originate in the House of Representatives or a lower house); (d) incarcerate citizens without showing just cause (Article I, Section 9.2 of the U.S. Constitution limits the suspension of the writ of habeas corpus); (e) charge persons with a crime without a grand jury trial (this is echoed in the Fifth Amendment of the U.S. Constitution); or (f) take away private property without providing just compensation (this is similarly found in the Fifth Amendment of the U.S. Constitution).

The most far-reaching limitations on monarchical rule, however, came with the English Bill of Rights. The Glorious Revolution of 1688, a bloodless revolt by Parliament against the monarchy, culminated in 1689 when the newly crowned royalty, William III and Mary, agreed to recognize 13 specified civil rights. Those enumerated rights, which bear striking resemblance to ones found in the U.S. Constitution, include the right of British subjects (a) to petition the monarch for a redress of grievances (a similar provision is found in the First Amendment of the U.S. Constitution); (b) to be protected against excessive bail and fines and cruel and unusual punishment (guarantees also found in the Eighth Amendment of the U.S. Constitution); (c) to a fair and speedy trial (provisions likewise found in the Sixth Amendment of the U.S. Constitution); (d) to free speech and debate in Parliament (a similar guarantee is afforded members of Congress in Article I, Section 6.1 of the U.S. Constitution); (e) to have Parliament approve the keeping of a standing army (Article I, Section 8.12 of the U.S. Constitution gives Congress financial control over the armed forces); and (f) the right of certain subjects (Protestants) to keep and bear arms (a similar provision is found in the Second Amendment of the U.S. Constitution).

Influence of Natural Rights Philosophers

A critical element of the British and American political experiences was the influence of the so-called natural rights philosophy that evolved during the late 17th and early 18th centuries, particularly the theories advanced by Thomas Hobbes, John Locke, and Jean-Jacques Rousseau. The common belief among these thinkers was that humans once lived in a state of nature where they had complete, unbridled freedom. However, this state of nature was fraught with peril and uncertainty, because there was nothing in the state of nature to protect their lives, property, or liberty. As Hobbes asserted in *Leviathan*, life in the state of

nature was “solitary, poor, nasty, brutish and short” (1909–1914, para. 9). To survive, the people entered into a social contract or compact with the sovereign. In exchange for security, they agreed to relinquish certain freedoms but never their natural or unalienable rights. For Hobbes, the fundamental natural right was the right to life. For Locke, natural rights included freedoms of expression and, most important, ownership of property. And Rousseau believed the social contract required subjects to trade unrestrained liberties for civil liberties, or freedoms within the law. Each of these philosophers realized the necessity of balancing liberty with authority or security. The difference among theories, of course, was a question of degree.

The natural rights document that doubtless had the greatest impact on the American political experience was John Locke’s *Two Treatises on Government* (1689/1764). Written to justify the Glorious Revolution of 1688 and the overthrow of King James II of England, Locke’s disquisition sets forth many ideas that were read by American colonists and served as blueprints for the American Revolution. Among other things, Locke posits that people are created equal and possess the rights to liberty and property. Locke further maintains that humans are entitled to pursue their own potentials in life. And should the sovereign fail to protect these “unalienable rights,” Locke asserts, the people have the right to overthrow the sovereign. As will be seen, these phrases bear striking resemblance to those echoed by Thomas Jefferson in the American Declaration of Independence of 1776. The U.S. Constitution’s (1787) Fifth and Fourteenth Amendments also stipulate that no person shall be deprived of “life, liberty, or property, without due process.” Finally, the Ninth Amendment carries natural rights overtones, declaring, “The enumeration in this Constitution of certain rights shall not be construed to deny or disparage others retained by the people.”

Influence of the Liberal Arts and Age of Enlightenment

Our nation’s founders were no doubt influenced by their education. Most of the delegates who signed the Declaration of Independence and drafted the Constitution were relatively well educated. Virtually all first learned to read in the home, most received some formal education, and many attended college at a relatively young age. Approximately 30 of the 55 delegates to the Constitutional Convention were college graduates. The centerpiece of the 18th century college education was the liberal arts curriculum, which was generally composed of what was termed the *trivium* (grammar, rhetoric, and logic) and *quadrivium* (arithmetic, geometry, music, and astronomy). Thus, the emphasis was on acquiring intellectual and cultural knowledge, as opposed to learning technical skills or a trade.

This too was the advent of the Age of Enlightenment, a period when intellectuals throughout Europe and the American colonies yearned to comprehend human nature

and sought to improve life through social and political reforms. Among the leaders of this intellectual movement was Sir Francis Bacon, an English scholar known for his theory of scientific discovery. Bacon posited that it was possible to observe nature, discover universal laws, and apply them to ameliorate society's problems. Many of the Constitution's framers were avid readers and were quite familiar with Bacon and other great thinkers of the period, such as David Hume, Edmund Burke, and Baron de Montesquieu. Even George Washington, who possessed but 5 years of formal schooling, collected books and was familiar with the classics. Perhaps Jefferson said it best: "I cannot live without books" (Cappon, 1987, p. 443).

Influence of Classical Republicanism

The American founders were also influenced by classical republicanism: ancient Greek and Roman political philosophy. The central tenet of classical republicanism is that governmental leaders should put the public interest above private desires and promote the common good through civic virtue. Civic virtue entails such ideals as individual accountability, responsibility, integrity, compassion, fairness, generosity, courage, self-control, and moderation. Records reveal that many of the leaders who forged the Declaration of Independence and drafted the Constitution were quite familiar with ancient Greek and Roman culture and history. Among the ancient Greek thinkers and orators they admired were Socrates, Plato, Aristotle, Sophocles, Plutarch, Thucydides, Polybius, Demosthenes, and Homer. The framers appeared equally enamored of the writings of Cicero, Horace, Justinian, and Virgil from the Roman Republic (circa 500 to 27 BCE). This admiration is well demonstrated by the fact that both anti-Federalist and Federalist pamphlets (see the following discussion) were published under the pseudonyms of Roman heroes, such as Brutus, Centinel, Cato, and Publius. Personal correspondence among the American founders, particularly Thomas Jefferson, John Adams, Alexander Hamilton, and James Madison, also reveals frequent references to these early Greek and Roman thinkers. And on relinquishing the presidency, George Washington was often referred to as America's Cincinnatus—a popular Roman consul who voluntarily gave up power to return to the plow.

Influence of Judeo-Christian Teachings

Finally, there is clear evidence our early American leaders were profoundly influenced by Judeo-Christian teachings. At least 9 of the original 13 colonies had established state churches prior to the Declaration of Independence. Many of the 18th-century leaders learned to read from the Bible and were versed in the scriptures. Colonial colleges often required students to translate Biblical passages into Latin and Greek as a condition of admission. Delegates to the Constitutional Convention frequently quoted the Ten Commandments, the Sermon on the Mount, and other

passages from the Old and New Testaments. Early state constitutions contained references to the Deity. The Declaration of Independence speaks of Nature's God and rights endowed by the Creator. The Liberty Bell's inscription, "Proclaim Liberty throughout the Land," was a passage from the book of Leviticus in the Old Testament. And Benjamin Franklin (1787) observed, "God governs in the affairs of men" (para. 2).

Movement Toward American Independence

Salutary Neglect

Nearly 170 years transpired between the founding of Jamestown and American independence. During that span, Great Britain and the American colonies maintained a symbiotic relationship. As noted above, the colonies supplied the mother country with raw materials and provided a market for goods manufactured in England. The colonies benefited from a common language, uniform currency, and military protection from the French, Spanish, Native Americans, and pirates on the high seas. But the 3,000 miles separating the colonies from the mother country created communication problems. It became impossible for the Crown to maintain tight control over the daily affairs in the colonies. Over time, the colonists gained considerable freedom from British regulations and frequently found ways to evade what were relatively low taxes. The British, of course, realized what was happening but largely turned a blind eye as long as the system was functioning. This situation became known as *salutary neglect*.

However, this rather tranquil relationship between Great Britain and the American colonies began to turn sour in 1754 with the outbreak of the French and Indian War (also called the Seven Years' War). In that year, intense fighting broke out between two coalitions over control of the North American continent. One faction was composed of the French and their Native American allies, and the other faction consisted of the British, American colonists, and their Native American supporters. The British-American-Native American coalition ultimately prevailed in 1763, with Great Britain wresting control of Canada, Louisiana, and the Mississippi River Valley.

This war proved particularly costly for Great Britain for two reasons. First, in 1754, the British convened a meeting of colonial representatives in Albany, New York, to organize for war. But the assemblage went far beyond British expectations. At this meeting, Benjamin Franklin of Pennsylvania introduced the Albany Plan of Union, calling for a colonial confederation to levy taxes, establish a militia, and address common concerns. The British rejected the plan, fearing it would weaken their control over the colonies, but the seeds were planted for unification. Second, the war significantly drained the British treasury. The cost of maintaining British troops in North America

while simultaneously fighting the French on the European continent was staggering. To pay for the war, the Parliament began to impose a series of taxes and trade regulations on the American colonists.

The American colonists greatly resented these new taxes and regulations. Among the most despised British taxes was the stamp tax of 1765, a tax on all printed materials, including playing cards, legal documents, diplomas, calendars, and newspapers. The colonists, believing this was taxation without representation, called for a meeting in New York City in October 1765 to voice their opposition. Nine colonies sent representatives to the Stamp Act Congress. The delegates petitioned the Crown and Parliament to rescind the new taxes and called for a boycott of British goods. To enforce this boycott, colonists formed groups called Sons and Daughters of Liberty. Among the most notable protests was the Boston Tea Party on December 16, 1773, when some 200 patriots, disguised as Indians, stormed three ships to throw tea into the harbor in protest of the British tea tax. The British reacted by imposing additional taxes in 1774, known as the Coercive or Intolerable Acts, which in turn galvanized the colonists' resolve.

First Continental Congress

To exert greater pressure on the Parliament, the 56 delegates from all states but Georgia assembled at Carpenter's Hall in Philadelphia on September 5, 1774. Notable delegates included Patrick Henry and George Washington of Virginia, John Adams and Samuel Adams of Massachusetts, John Dickenson of Pennsylvania, John Rutledge of South Carolina, and Roger Sherman of Connecticut. Known as the First Continental Congress, the delegates denounced the Intolerable Acts, petitioned King George III to make reforms, and agreed to meet the following year should their demands be ignored. The British again refused to accede to the colonists' demands. Then, on April 19, 1775, fighting broke out at Lexington and Concord in Massachusetts between the British regulars and the colonial militia. These so-called shots heard round the world signaled the start of the Revolutionary War.

Second Continental Congress

The delegates reassembled at the State House (now called Independence Hall) in Philadelphia on May 10, 1775, becoming the Second Continental Congress. This time, all 13 colonies sent representatives. Joining this congress were John Hancock and Elbridge Gerry of Massachusetts; Oliver Ellsworth of Connecticut; John Jay, Robert Livingston, and Alexander Hamilton of New York; James Madison, George Wythe, Edmund Randolph, and Richard Henry Lee of Virginia; Benjamin Franklin and James Wilson of Pennsylvania; and Charles Pinckney of South Carolina. Initially, the delegates were badly split on the

issue of breaking ties with Great Britain. Roughly one third of the delegates opposed independence, one third favored independence, and one third were uncertain how to proceed. John Hancock of Massachusetts was chosen the president of the Continental Congress, as the representative voted to raise money, issue currency, and form a continental army under the leadership of George Washington.

As events unfurled, public support for independence accelerated. Among the most persuasive arguments for independence were those advanced in Thomas Paine's *Common Sense*. This widely read pamphlet, appearing in January 1776, underscored what Paine considered the people's moral obligation to cast off the yoke of British oppression. Within months, calls for independence resonated throughout the colonies. Then, on June 7, 1776, Richard Henry Lee introduced this bold motion to the Second Continental Congress:

Resolved, That these United Colonies are, and of right ought to be free and independent States, that they are absolved from all allegiance to the British Crown, and that all political connection between them and the State of Great Britain is, and ought to be, totally dissolved. (para. 1)

The congress approved the resolution, and a committee, composed of Robert Livingston, Benjamin Franklin, John Adams, Roger Sherman, and Thomas Jefferson, was appointed to draft the formal declaration. The bulk of this document, the Declaration of Independence, was written by the erudite Jefferson.

The Declaration of Independence

The Declaration of Independence, often called the birth certificate of our nation, can be viewed as three interrelated sections. In the first section, Jefferson explained the reason for the declaration and the philosophical principles on which it is constructed. He began by stating the underlying principles of democracy—representative government, limited government, rule by law, and individual democracy. Borrowing extensively from the natural rights philosophers and the Judeo-Christian heritage, Jefferson wrote,

We hold these truths to be self evident, that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty and the pursuit of Happiness. That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed. (Declaration of Independence, para. 3)

Note how Jefferson substituted the words *pursuit of happiness* for the Lockean concept of property. Some scholars believe the phrase *pursuit of happiness* was just a form of shorthand for property, while others maintain the substitution was effectuated to prevent Americans from falsely anticipating the distribution or redistribution of property by

the new government. Again, note the Constitution's Fifth and Fourteenth Amendments' reinsertion of property when they guarantee that no person shall be denied the right to "life, liberty or property, without due process."

The second section of the Declaration provides a detailed list of grievances against King George III. Among other things, the document charges that he has quartered British soldiers in civilian homes, levied taxes without the consent of the people, cut off trade, denied the right to trial by jury, abolished elected legislatures, and imposed absolute rule.

The final section clarifies the colonists' fruitless efforts to resolve differences peacefully, chastises the Crown's recalcitrance, and declares that the colonists must be free of British rule. The document concludes with these searing words:

That these United Colonies are, and of Right ought to be Free and Independent States; that they are Absolved from all Allegiance to the British Crown, and that all political connection between them and the State of Great Britain, is and ought to be totally dissolved; and that as Free and Independent States, they have full Power to levy War, conclude Peace, contract alliances, establish Commerce, and to do all other Acts and Things which Independent States may of right do. (Declaration of Independence, para. 33)

Technically, the Second Continental Congress formally declared independence on July 2, but it would be another 2 days before the redacted Declaration of Independence received final approval on July 4, 1776. The document was then delivered to John Dunlap, a renowned Philadelphia artisan, for final printing. Contrary to popular paintings, the delegates did not sign the Declaration of Independence at one sitting. Rather, the 56 signatures were affixed to the engrossed copy over a period of weeks, beginning with John Hancock's bold strokes on August 2, 1776. Although the Declaration of Independence makes it clear that the colonists were cutting ties with England, uncertainty remained about the nature of the new government. Were the signers creating 13 sovereign nations or 1 sovereign nation consisting of 13 states? The document appeared to send mixed signals by stating that "these United Colonies are . . . Free and Independent States" (Declaration of Independence, para. 33).

Colonies Transform Into States

Clearly, the colonies thought they were independent states. The process of transforming colonies into independent states started even before the Declaration of Independence was inked. New Hampshire became the first colony to jettison its royal charter and craft a state constitution in January of 1776, and over the course of several years, the remaining colonies followed suit. Each new state considered itself an independent, sovereign nation, free to engage in foreign commerce, negotiate treaties, establish

currency, and control its own borders. Although the 13 new state constitutions varied in wording and length, they contained a number of commonalities. All of them embraced separation of powers, establishing a legislative, an executive, and a judicial branch of government. With the exceptions of Pennsylvania and Georgia, all adopted bicameral or two-house legislatures, with the lower house being popularly elected by qualified voters. For the most part, the executive branch was considered weak, reflecting the recent colonial distrust of the royal governors. Seven states also included a bill of rights. And each newly created state continued to send delegates to the Second Continental Congress and support the establishment of the Articles of Confederation.

Articles of Confederation

As the Revolutionary War accelerated, representatives at the Second Continental Congress soon came to the realization that some central governmental apparatus was necessary to coordinate the states, marshal resources, and prosecute the war. In 1776, a committee set forth a plan for the Articles of Confederation. The proposal was debated for more than a year and a half before the Continental Congress approved it and it was sent to the states for ratification. The Articles of Confederation (1777) went into effect in 1781, after Maryland became the requisite 13th state to ratify the plan. The Articles thus became America's first written national constitution.

The Articles of Confederation formed largely an association of independent states. Its stated purpose was to establish a "firm league of friendship" to promote "their common defense, the security of their liberties, and their mutual and general welfare" (Article III). Yet the document recognized each state's right to retain its "sovereignty, freedom, and independence" (Article II). The centerpiece of the Articles was a unicameral legislature. All representatives were appointed by their respective state legislatures, and most received only nominal compensation at best. The number of representatives allotted for each state delegation ranged from two to seven, but each state was entitled to cast a single vote in the Confederation Congress. All legislative actions required approval of two thirds, or 9 of the 13 states. There was no executive authority, and the judiciary had extremely limited jurisdiction. Thus, the central government was inherently weak.

From the very outset, the Articles of Confederation were plagued by a series of vexing problems. The Confederation Congress had no authority to impose tariffs or collect taxes. Without revenue, the confederation could not adequately pay an army or a navy. And without military forces, the confederation could not prevent piracy on the high seas, protect settlements on the frontier, or pay off a growing war debt. The Confederation Congress was also powerless to regulate commerce or settle boundary disputes between sister states. States frequently taxed imported goods,

coined their own money, imposed trade barriers, and negotiated treaties with foreign governments and Native American tribes. Because states often refused to recognize contracts, civil judgments, and criminal proceedings from neighboring states, debtors could stave off bill collectors and criminals could avoid prosecution by simply moving across state lines. Finally, it was virtually impossible to address these glaring weaknesses because unanimity was required to amend the Articles.

The Articles of Confederation, however, were not a total failure. Despite their glaring weakness, they managed to hold the states together long enough to conclude the war and forge the Treaty of Paris with Great Britain in 1783. The Articles also enacted the Northwest Ordinance of 1787 that provided the ground rules for developing Western lands, annexing territories, and admitting new states under the U.S. Constitution (Article IV, Section 3.1 and 3.2). The Articles diplomatic corps also provided international experience for many future American leaders, including Thomas Jefferson, John Adams, and Benjamin Franklin. And some of the legislative committees and administrative rules established under the Articles laid the foundation for our present-day Departments of State, Treasury, and Defense.

The Constitution of the United States

Road to the Constitution

By 1786, the Confederation was on the verge of economic collapse. The currency issued by the Continental Congress became seriously devalued, and states accelerated the printing of worthless paper money to enable citizens to pay their debts and taxes. Inflation was running rampant, and foreclosures reached epidemic proportions. Citizens gathered in the streets and at government buildings to protest and advocate for government reform. Emblematic of the growing economic unrest was Shay's Rebellion in Massachusetts. In 1786, Daniel Shay, a distraught farmer and Revolutionary War veteran, organized protests of debtors' prisons and courthouses. The uprising was ultimately quelled, but it sent a clear signal among state leaders that drastic action was needed to restore public confidence and restore order.

A group of prominent Virginians, led by George Washington and James Madison, convinced the Virginia legislature to convene a meeting of sister states in Annapolis, Maryland, to discuss maritime commerce, trade, and the state of the union. In September 1786, five states—Virginia, New York, Delaware, New Jersey, and Pennsylvania—sent representatives to this Annapolis Convention. Among the most influential representatives that met at Mann's Tavern were James Madison (Virginia), Alexander Hamilton (New York), John Dickenson (Delaware), and Edmond Randolph (Virginia). After discussing their common economic woes, the small assemblage

issued a challenge to all states to consider revising the Articles of Confederation. Spurred on by the Annapolis Convention, in February 1787, the Continental Congress officially called on all 13 states to send delegates to Philadelphia the following May:

for the sole and express purpose of revising the Articles of Confederation and reporting to Congress and the several legislatures such alterations and provisions therein as shall . . . render the federal constitution [the Articles of Confederation] adequate to the exigencies of Government and the preservation of the Union. (Resolution of Congress, 1787, para. 2)

The Constitutional Convention

The Framers

Eleven states quickly appointed delegates to assemble at the State House (now called Independence Hall) in Philadelphia on May 14, 1787. New Hampshire's delegates appeared much later in the proceedings, and Rhode Island opted not to participate. Altogether, 73 delegates were appointed, only 55 attended any of the sessions, and just 39 signed the completed document. All were white males, and all were Protestant, save one (Daniel Carroll, a Catholic from Maryland). The delegates' average age was roughly 43 years. Jonathan Dayton of New Jersey was the youngest delegate, at 26, and the venerable Benjamin Franklin was the oldest, at 81. Occupationally, 32 of the 55 delegates were lawyers, 8 were merchants, 3 were medical doctors, 2 were college presidents, about 2 dozen owned slaves, 1 was an inventor (Benjamin Franklin), and nearly all had considerable political seasoning. Eight had signed the Declaration of Independence, 44 were state representatives in the Continental Congress, 7 had been governors, many had been officers in the Continental army, and the vast majority served in their respective colonial or state legislatures.

Unquestionably, the most prominent national figures to participate in the Constitutional Convention were George Washington of Virginia and Benjamin Franklin of Pennsylvania; they provided a sense of legitimacy and gravitas to the convention. Other highly respected senior solons included Roger Sherman of Connecticut, John Dickinson of Delaware, George Mason of Virginia, John Rutledge of South Carolina, and William Livingston of New Jersey. Among those who held strong state and parochial interests were William Patterson of New Jersey, Charles Pinckney of South Carolina, Luther Martin of Maryland, and Robert Yates and John Lansing of New York. Those advocating strong national proclivities included James Wilson of Pennsylvania, Alexander Hamilton of New York, and James Madison of Virginia. Noticeably absent were patriots Thomas Jefferson, John Adams, John Hancock, Samuel Adams, and Patrick Henry. Jefferson and John Adams were serving diplomatic missions in Europe, Hancock and Sam Adams were

not picked as their respective states' delegates, and Patrick Henry held apprehensions about the convention's motives, saying he "smelt a rat."

The Convention Organization

The Convention did not reach a quorum (7 of 12 states) to conduct a formal meeting until May 25, 1787. The first order of business was to select its leaders and establish rules of procedure. General George Washington was unanimously elected the presiding officer, and Major William Jackson of South Carolina was selected recording secretary. Washington sat at the front of the assemblage and appeared to say little during the proceedings, but his towering presence and air of neutrality doubtlessly helped maintain decorum. Jackson's official journals, however, were slipshod and incomplete. Fortunately for posterity, others took notes, including Robert Yates, James McHenry, John Lansing, and, most important, James Madison. Madison sat near the front of the convention and maintained the most detailed records of the convention's motions and debates. It is largely for this reason, plus his sage proposals, that James Madison has been appropriately dubbed the Father of the Constitution. (Madison's notes were subsequently given to Washington, who deposited them with the Department of State in 1796. Following Madison's death in 1836, Congress promulgated the convention notes.)

The convention rules were quite simple and flexible. The framers agreed to keep the proceedings secret to encourage candor and full discussion of the issues. Doors remained locked and guarded. And to maintain secrecy, the delegates often had to conduct business with the windows down during the long, hot summer months. Much of the proceedings took place in the committee of the whole, and specific proposals were assigned to select committees to help iron out differences. Each state had but one vote, a quorum was required, and motions passed with a simple majority vote.

The Challenges

The challenges facing the convention delegates were both daunting and complex. James Madison would later summarize the central dilemma this way: "In framing a government which is to be administered by men over men, the chief difficulty lies in this: You must first enable the government to control the governed, and in the next place oblige it to control itself" ("The Federalist No. 51," 1788, para. 4). Obviously, the framers realized the Articles of Confederation had failed to adequately control the government and needed to be jettisoned. But they were equally fearful of establishing a central government so powerful that it would run roughshod over the states and, more important, the people's rights. Was it possible to create a written constitution that could satisfy both

objectives? This is something that no national polity had theretofore ever created. And how would the framers address the critical issues regarding representation, slavery, and the economy?

Forging the New Constitution

The Issue of Representation

Once the Convention had agreed on its rules, Edmund Randolph of Virginia quickly introduced a groundbreaking plan that framed the debate over the new constitution. This plan, known as the *Virginia Plan*, was really the brainchild of James Madison. It contained 15 resolutions that called for the abandonment of the Articles of Confederation and the creation of a strong national government. Among other things, the plan called for three branches of government—legislative, executive, and judicial. The legislative branch would be composed of two houses, an upper and a lower house, whose representation would be based on a state's population. Virginia, for example, would have 16 representatives, while Delaware would be entitled to just two. The popularly elected lower house would then select members of the upper house. The legislative branch would also choose members of the executive and judicial branches, but each branch would possess the ability to check the other branch. Additionally, the national government would possess authority to override contradictory state laws. The convention debated this plan for nearly a fortnight. The more populous states, such as Massachusetts, New York, and Virginia, naturally favored this large-state plan, while the less populous states, such as Delaware, Georgia, and New Jersey, understandably vehemently opposed the Virginia Plan.

On June 15, William Patterson of New Jersey offered a counterproposal, known as the small-state or *New Jersey Plan*. Under this plan, the Articles of Confederation essentially would be maintained but significantly altered. The plan called for a one-house congress with representatives appointed by their respective state legislatures. And each state would have but one vote, irrespective of the state's population. However, the plan called for plural executives chosen by the national congress and a supreme court appointed by the executive branch. The convention debated this plan for 4 days, but its provisions, and alternatives offered by Alexander Hamilton of New York and Charles Pinckney of South Carolina, were categorically rejected by the larger states. The delegates were thus at loggerheads over the issue of representation.

By late June, the convention appeared to be on the verge of collapsing over the issue of representation and structure of the national government. Indeed, Luther Martin (1911), of Maryland, recorded on June 28 that the convention was "scarcely held together by the strength of a hair" (p. 196). The stalemate was finally broken when

the Connecticut delegation, led by Roger Sherman, offered this solution: “In one branch the people ought to be represented; in the other, the states” (Boutell, 1896, p. 145). According to this Connecticut Compromise (or Great Compromise), the new congress would be composed of two houses—a lower and an upper house. Representatives in the lower house, or House of Representatives, would be apportioned according to population and would be elected by qualified voters in their respective states. This chamber would also be charged with initiating revenue bills. Legislators in the upper house, or Senate, would be chosen by their respective state legislatures, and every state would be entitled to the same number of senators.

The Issue of Slavery

The framers also faced the vexing issues relating to slavery. The total population among the 13 original states (excluding Native Americans) was approximately 3 million in 1787, and of that, roughly 300,000, or about 10% of the total population, were slaves. The vast majority of slaves lived in the six Southern states, where slave labor was critical to supporting large plantations. Delegates from the Southern states were therefore concerned over four critical issues. First, they opposed any efforts to abolish or tax the importation of slaves under the new constitution. Second, Southern delegates sought assurances that free states would return runaway slaves to their owners. Third, Southern delegates wanted slaves to count for representation in the national legislature. But fourth, they did not want slaves counted for the purpose of apportioning taxes among the states. Northern delegates naturally opposed such measures; they believed Southern states would accelerate the importation of slaves to unfairly increase their representation in the lower house.

This sectional rift over slavery too was resolved through a series of compromises. Regarding the issue of importation, the delegates agreed, in the so-called slave trade compromise, that the importation of slaves “shall not be prohibited by the Congress prior to the year one thousand eight hundred and eight” and that an importation tax shall not exceed “ten dollars for each Person” (U.S. Constitution, 1787, Article I, Section 9.1). (Congress outlawed the importation of new slaves in 1808, and slavery, of course, was ended with the Civil War and the passage of the Thirteenth Amendment in 1866.) The framers also agreed to the so-called fugitive slave compromise in Article IV, Section 2.3, authorizing Congress to establish rules for recapturing escaped slaves. (This was rendered null and void with the passage of the Thirteenth Amendment.) Finally, concerning the issue of representation and taxation, the delegates agreed in Article I, Section 2.3, to a so-called three-fifths compromise, whereby slaves were to be counted as three fifths of free persons. (This too was superseded by Section 2 of the Fourteenth Amendment in 1868.)

Economic Issues

Delegates were also badly divided over economic issues. The Northern states’ delegates were particularly concerned about commerce. As noted earlier, Northern manufacturers and shipping interests were hard hit under the Articles of Confederation. States frequently imposed trade barriers, coined their own money, and were unable to enforce contracts made in sister states. The national government was also powerless to regulate commerce, protect domestic goods from foreign competition, or prevent the piracy of manufactured goods shipped on the high seas. Northern delegates thus preferred a strong national government with the power to negotiate treaties, levy taxes, and regulate commerce. Southern states’ delegates, however, feared that a strong national government might run roughshod over their economic interests. More specifically, they feared a strong national government would impose tariffs that might have deleterious effects on their agricultural products (namely, tobacco and cotton) and engage in treaties that would ultimately outlaw slave transactions.

Concessions were made on both sides of the economic divide. Among other things, the Northern states’ delegates were able to secure the national governmental power to (a) lay and collect taxes (Article I, Section 8.1), (b) negotiate treaties (Article II, Section 2.2), (c) regulate foreign and interstate commerce (Article I, Section 8.3), (d) provide for an army and navy (Article I, Sections 8.12 and 8.13), and (e) punish piracy on the high seas (Article I, Section 8.10). The Southern states’ delegates won assurances that (a) Congress could not tax articles exported from states (Article I, Section 9.5), (b) Congress could not prefer one port over another (Article I, Section 9.6), (c) treaties required a two-thirds vote of the Senate (Article II, Section 2.2), and (d) the slave trade would not be prohibited for two decades (Article I, Section 9.1).

Basic Features of the Constitution

Separation of Powers

Among the most striking features of the U.S. Constitution is the concept of separation of powers. Separation of powers can be traced to French philosopher Baron de Montesquieu (1748/1989) and his classic work, *The Spirit of the Laws*. Drawing on the ancient Roman republic, Montesquieu believed there were three fundamental forms of government—monarchy (rule by one), aristocracy (rule by few), and democracy (rule by many)—and that the best way to achieve the common good was to have a mixed constitution that combined all three forms. Montesquieu further believed the British government of his day more or less encapsulated this notion with the tripartite division of the king (monarchy), the House of Lords (aristocracy), and the House of Commons (democracy).

The framers, particularly Madison, were familiar with Montesquieu’s (1748/1989) concept of dividing powers

and the necessity of preventing any one interest from becoming too powerful. However, the framers modified Montesquieu's concept by creating three separate branches of government in the U.S. Constitution. Article I, Section 1 states, "All Legislative Powers herein granted shall be vested in a Congress which shall consist of a Senate and a House of Representatives." Article II, Section 1 provides, "The Executive Powers shall be vested in a President of the United States. . . ." And Article III, Section 1 begins, "The Judicial Power shall be vested in one Supreme Court and such inferior courts as Congress shall from time to time ordain and establish."

Under this plan, each of the three branches is responsible to a different constituency. Members of the Senate were to be selected by their respective state legislatures (the Seventeenth Amendment in 1913 calls for the popular election of U.S. senators), while members of the House of Representatives were to be popularly elected by qualified voters in their respective states. The president of the United States is to be chosen by electors (collectively referred to as the electoral college). And the Supreme Court justices are appointed by the president with "advice and consent" (a simple majority vote) of the Senate. Additionally, the officials of each branch serve for different terms, making it difficult for any one interest to wrest control of government in a single election. Thus, U.S. senators serve 6-year staggered terms so that only one third are up for reelection every 2 years, U.S. House members serve 2-year terms with all members facing reelection every 2 years, the president serves 4-year terms (the 22nd Amendment of 1951 restricts the president to two 4-year terms or a total of 10 years), and justices of the Supreme Court serve as long as they maintain "good behavior."

Checks and Balances

In addition to separating the powers, each branch is given constitutional authority to challenge the other two branches. This makes it difficult for any particular interest to gain too much control of the national government. This arrangement, known as checks and balances, also comports with Madison's belief that "ambition must be made to counter ambition" ("The Federalist No. 51," 1788, para. 4). Consider some of the powers and restrictions the framers built into our Constitution. A bill that passes the House of Representatives can be tabled in the Senate, and vice versa (Article I, Section 7.2). Any bill that passes both houses of Congress can be vetoed by the president (Article I, Section 7.2), but vetoes (excluding pocket veto) can be overridden by a two-thirds vote of both houses of Congress (Article I, Section 7.3). The Supreme Court can declare acts of the Congress unconstitutional (see judicial review in the following discussion), but Congress can alter the appellate jurisdiction of the federal courts (Article III, Section 1) and propose constitutional amendments (Article V) to supersede the high court (as it did with the Eleventh,

Fourteenth, and Sixteenth Amendments). And the House can impeach and the Senate can remove the president or federal judges for "Treason, Bribery, or other high crimes and Misdemeanors" (Article I, Section 3.6; Article II, Section 4).

Federalism

In addition to the issues of representation and slavery, the question of how to divide powers between the national government and the states was of paramount concern to the framers. As noted earlier, the delegates had personally experienced the weaknesses of a confederation and recognized the necessity of establishing a stronger national government. Yet delegates were leery of a unitary government and were unwilling to surrender their respective states' control over local affairs. The compromise was to establish a totally new and unique form of government, a federal form of government, in which a written constitution enumerates the powers belonging to the national government and places restrictions on both the national and state governments.

More specifically, the original (sans amendments) U.S. Constitution (a) delegates powers to the national government—Congress, for example, has the power to lay and collect taxes and tariffs (Article I, Section 8.1), declare war (Article I, Section 8.11), promote science and arts (Article I, Section 8.8), suppress insurrections and repel invasions (Article I, Section 8.15), create new states (Article IV, Section 3), and guarantee each state a republican form of government (Article IV, Section 4); (b) restricts powers of the national government—Congress may not suspend the writ of habeas corpus in peacetime (Article I, Section 9.2), pass ex post facto laws or bills of attainder (Article I, Section 9.3), tax goods exported from states (Article I, Section 9.5), or grant titles of nobility (Article I, Section 9.8); and (c) restricts the powers of the state governments—state governments, for instance, may not enter into treaties with foreign governments, grant letters of marque and reprisal (authority of private bounty hunters to capture pirates), coin money, pass bills of attainder or ex post facto laws, interfere with private contracts, or grant titles of nobility (Article I, Section 10.1).

Additional governmental delineations were created after the U.S. Constitution was ratified. Among the powers recognized through subsequent constitutional amendments or Supreme Court rulings were (a) powers retained by the people—see the Ninth Amendment and rulings on the right of marital privacy (*Griswold v. Connecticut*, 1965) and abortion rights (*Roe v. Wade*, 1973); (b) powers reserved to the state or people—powers that are neither given to the national government nor prohibited the states (see the Tenth Amendment) and include the state governments' right to provide public education or regulate political parties (because neither terms are found in the U.S. Constitution); and (c) powers implied from the Constitution—these powers

are not specifically stated in the Constitution but are naturally inferred from those that are. Implied powers can be traced to the landmark decision *McCulloch v. Maryland* (1819), when Chief Justice John Marshall upheld the federal government's authority to establish a national banking system, even though the word *bank* nowhere appears in the constitution. Marshall reasoned, in part, that since Congress possessed the delegated powers to coin money, borrow money, tax, spend, punish counterfeiting, and regulate bankruptcies (see Article I, Section 8) and since Congress is also empowered to do what is "necessary and proper" (see Article I, Section 8.18, the so-called elastic clause) to carry out those delegated powers, Congress therefore had the right to establish a banking system. The Constitution also placed limits on government, including (a) restrictions on the national government—among the rights protected were freedom of expression (First Amendment), the right to bear arms (Second Amendment), freedom from unreasonable searches and seizures (Fourth Amendment), freedom from self-incrimination (Fifth Amendment), the right to counsel (Sixth Amendment), and freedom from cruel or unusual punishments (Eighth Amendment); and (b) restrictions on the state governments—the U.S. Constitution now proscribes state governments from denying persons the right to due process, equal protection, privileges, or immunities (see the Fourteenth Amendment). These are just a few of the proscriptions set forth in the Constitution and Bill of Rights and later amendments.

Supremacy of National Laws

The framers understood that there would be periodic conflicts between the national laws and state laws, especially over concurrent or shared powers (for example, both national and state governments may levy taxes, charter corporations, and enact criminal laws). To resolve jurisdictional conflicts between the two levels of government, the founders included this so-called supremacy clause:

This Constitution, and the law of the United States which shall be made in pursuance thereof; and all treaties made, or which shall be made, under the authority of the United States, shall be the supreme law of the land; and the judges in every state shall be bound thereby, anything in the Constitution or laws of any state to the contrary notwithstanding. (U.S. Constitution, 1787, Article VI, Section 2)

Thus, if a conflict occurs between the U.S. Supreme Court ruling and a state constitution, the former would hold sway (see *Brown v. Board of Education*, 1954). Similarly, should a dispute arise over a legitimate treaty and a state constitution, the treaty would take precedence (see *Missouri v. Holland*, 1920). And if a dispute arises between a law passed by Congress and a state legislature, the national statute generally would be deemed superior (for example, see *Pennsylvania v. Nelson*, 1956). The supremacy clause is the legal mortar that helps hold the federal system together.

Ratifying the Constitution

The framers drafted, debated, and concluded their work in 116 days. On September 17, 1787, 39 brave delegates, starting with George Washington, affixed their signatures to the final document. The Constitution of the United States is a relatively short document providing the framework on which the American political system is built. The final version was penned by Jacob Shallus, a clerk for the Pennsylvania state legislature, for a fee of \$30, on four pages of parchment, approximately 24 inches wide and 29 inches long. The document contains just 4,400 words (perhaps fewer words than today's daily newspapers' sports sections). It begins with the Preamble, a one-paragraph introduction, stating its lofty purpose:

We the People of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defence [sic], promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America. (U.S. Constitution, 1787, para. 1)

The body of the Constitution is composed of the following Articles or legal subdivisions: Article I (legislative branch); Article II (executive branch); Article III (judicial branch); Article IV (interstate relations); Article V (amendment process); Article VI (debt assumption, supremacy clause, and oath restrictions); and Article VII (provisions for ratification).

Drafting the new document was a monumental accomplishment, but it was just the first step in the tortuous process. Next, the framers reported back to the Continental Congress, which reluctantly agreed to submit the proposed Constitution to the states for ratification on September 28, 1787. The Congress agreed to follow the procedures set forth in Article VII by requiring the approval of 9 of 13 states in ratifying conventions. The framers wisely opted to use state ratifying conventions rather than state legislatures, because many state legislators opposed efforts to strengthen the national government, especially in New York, Virginia, and Rhode Island (which, of course, sent no delegates to the Philadelphia convention).

Delaware became the first state to ratify the new Constitution on December 7, 1787, and over the course of 7 months, Pennsylvania, New Jersey, Georgia, Connecticut, Massachusetts, Maryland, and South Carolina followed suit. On June 21, 1788, New Hampshire became the ninth state to complete the requisite ratification. Still, there was considerable work to be done. Virginia and New York, the most populous and influential states, had yet to ratify the Constitution, and without their critical support, many feared the new government might not survive. For the most part, civic leaders split into two diametrically opposed camps—the anti-Federalists and the Federalists.

Anti-Federalists and Federalists Debate

Anti Federalists

Even before the Constitution's ink was dry, opposition to its provisions began to organize. Among the delegates who refused to sign the Constitution were Elbridge Gerry of Massachusetts and George Mason of Virginia. Other notable state leaders who joined the anti-Federalist camp included Patrick Henry, William Grayson, James Monroe, and Richard Henry Lee of Virginia; Samuel Adams and John Hancock of Massachusetts; Luther Martin of Maryland; Samuel Bryan of Pennsylvania; and George Clinton and Robert Yates of New York. During the ratification period, some anti-Federalists delivered speeches and published newspaper articles criticizing the new government. Those articles, now collectively known as the "Anti-Federalist Papers," appeared under such nom de plumes as Centinel (Samuel Bryan), Cato (George Clinton), Federal Farmer (most likely Richard Henry Lee), and Brutus (Robert Yates).

The anti-Federalists registered many complaints about the Constitution. Many objected to the secrecy enshrouding the Philadelphia convention. However, the most salient anti-Federalist objections concerned expanded powers of the national government relative to the states and citizens. They feared that the so-called necessary and proper clause (Article I, Section 8.18) would give Congress a virtual blank check to expand legislative authority over the states. They averred the president would have too many powers, especially military command of a permanent national army during peacetime. The prospect that the Supreme Court could nullify state constitutions and statutes using the so-called supremacy clause (Article VI, Section 2) was especially troubling. Most important, they complained the new constitution lacked a bill of rights to ensure individual liberties against the national government's action. The unamended Constitution, for example, contained no guarantees of free speech, press, or assembly.

The Federalists

Noteworthy supporters of the new Constitution included George Washington, William Randolph, John Marshall, and James Madison of Virginia; Gouverneur Morris of Pennsylvania; and Alexander Hamilton and John Jay of New York. Ironically, at the Philadelphia Convention, Alexander Hamilton had serious misgivings about the Constitution but signed begrudgingly, while William Randolph refused to sign altogether. The two, however, came to realize the futility of the Articles of Confederation and the necessity of embracing a stronger national government.

The most effective advocates of the proposed Constitution were Madison, Hamilton, and Jay. Between October 2, 1787, and May 28, 1788, the three published

85 essays in New York newspapers under the pen name Publius. Known collectively as "The Federalist Papers," the essays were widely read and effectively countered the arguments advanced by the anti-Federalists. For example, the Federalists maintained that the Articles of Confederation were beyond repair, that the separation of powers with built-in checks and balances would prevent any one branch from dominating the national government, that the popularly elected chamber (House of Representatives) would control the purse strings of military operations, that the Supreme Court would be the weakest branch of government, and that individual freedoms were protected under Article I, Sections 9 and 10. But to allay the anti-Federalist's fears, the Federalists promised to add a bill of rights once the Constitution had been ratified, and a national government was in place.

The Federalists' pledge to add a bill of rights doubtlessly helped sway members of the Virginia and New York ratifying conventions. Virginia ratified the Constitution on June 25, 1788, and New York ratified by a scant three votes on July 26, 1788. Ratification by these two critical states ensured that the new government would take root. The final business of the Continental Congress was to determine a temporary nation's capital, call for congressional elections, and set the date for the new government to commence.

The Constitution: An Elite Document or a Bundle of Compromises?

Over the years, many scholars have devoted considerable attention to ascertaining the motives of the people who wrote and supported the Constitution of the United States. The central question is whether the founders were motivated by self-interests or lofty principles. There are loosely two schools of thought—those who contend the Constitution reflects the preferences of an economic elite class, and those who believe the Constitution is not only an amalgam of compromised interests but also an embodiment of democratic theory.

Arguably the most influential scholar on this topic is Charles A. Beard (1913), a historian from the University of Wisconsin and a former president of the American Political Science Association. In his seminal work, *An Economic Interpretation of the Constitution*, Beard impugned the notion that the 55 delegates who attended the Constitutional Convention did so solely for altruistic reasons. Beard's rather cursory examination of treasury, census, and tax records led him to conclude that the framers were economically motivated to form the new government. He implies, then, that the U.S. Constitution is essentially an economic document created by political elites to protect their investments that were deleteriously impacted by the ineffectual Articles of Confederation. His research revealed that of the 55 convention delegates,

11 could be classified as manufacturers, merchants, or shippers; 40 were holders of public security interest (including George Washington); 14 were Western land speculators (including George Washington); 24 were lenders or creditors (including George Washington); 15 owned plantations with significant slaveholdings (including George Washington); but not a single delegate represented the so-called mechanics class (laborers). Beard further asserts that George Washington was likely the richest man in the 13 states and that the framers were primarily well-heeled lawyers from coastal regions where manufacturing and shipping were quite prevalent. Finally, Beard concludes that at least five sixths of the framers were the direct beneficiaries of the Constitution and that less than one sixth of the white male population had any voice in the ratification process.

Again, consider some of the provisions relating to economic interests in the U.S. Constitution. As noted earlier, Congress has the power to coin money, regulate interstate commerce, regulate bankruptcies, and establish weights and measures (Article I, Section 8). Such authority was lacking under the Articles. The U.S. Constitution grants the national government authority to establish an army (which can protect land investments on the frontier) and a navy (which can prevent piracy of shipped goods on the high seas). States are constitutionally forbidden (Article I, Section 10.1) to interfere with contracts (thus preventing debtor-dominated legislatures canceling money owed the creditors). The so-called full faith and credit provision (Article IV, Section 1) makes it difficult for people to avoid debts by moving to another state. People fleeing to another state to avoid criminal prosecutions may be extradited (Article IV, Section 2.2), and Congress regulates the development of Western lands (Article IV, Section 3.2). These and other economic-related clauses thus appear to lend credence to Beard's thesis.

However, there is a significant body of research that questions Beard's methodology and conclusion that the framers represented but a single, monolithic interest—the well-to-do. In his seminal work, *The Framing of the Constitution of the United States*, Max Farrand (1913) observed that the Constitution is a “bundle of compromises” of varied societal and economic interests. Similarly, noted historian Robert E. Brown (1956) asserted in *Charles A. Beard and the Constitution* that Beard seriously underplays the variety of economic interests at the Philadelphia Convention. Brown's empirical research reveals that poor people and middle-class Americans were also adversely affected under the Articles of Confederation, that the framers and state leaders who ratified the Constitution had democratic leanings, and that it is wrong to conclude the Constitution is an economic document created to protect the interests of wealthy creditors.

Arguably the most damaging criticism of Beard's thesis came from historian Forrest McDonald (1976), in his classic work, *We the People: The Economic Origins of the Constitution*.

McDonald began by reinvestigating the financial archives of the 55 delegates to the Philadelphia Convention plus the personal backgrounds of over 1,700 delegates to the state ratifying conventions. His perusal reveals that Beard's research was slipshod and that there were several dozen competing economic interests surrounding the adoption of the Constitution. Although McDonald concedes that most of the framers were relatively well-to-do, he also amply demonstrates there were many wealthy civic leaders, including those who held public securities, who seriously opposed the Constitution. Conversely, McDonald identified numerous people of modest means with scant property ownership who supported the Constitution. McDonald ultimately concludes that the process surrounding the adoption and ratification of the Constitution was fluid and dynamic and that the framers, as a whole, were principled, pragmatic men who understood the necessity of chartering a new form of government.

The New Government and The Bill of Rights

The Continental Congress chose Federal Hall on New York City's Wall Street to serve as the nation's first capitol and set the first Wednesday in February 1789 to elect members of the first Congress. Although the Constitution called for the Congress to convene on March 3, the new government lacked a quorum (simple majority) until April 6, 1789. Moreover, the new government commenced with just 11 state delegations. North Carolina did not ratify the U.S. Constitution until November 21, 1798, and Rhode Island became the 13th state to join the Union on May 29, 1790.

Among the new Congress's most pressing business was to select the first president, establish the judicial system, and draft a list of protected rights. George Washington was the unanimous choice of presidential electors (now popularly called the electoral college) and was inaugurated on April 30. With the passage of the Federal Judiciary Act of 1789, Congress set the number of justices on the Supreme Court at six (one chief justice and five associate justices) and determined that justices would ride circuit and decide cases twice yearly in the nation's capital. The Federal Judiciary Act also established the appellate jurisdiction of the Supreme Court (as opposed to the original jurisdiction that is found in Article III, Section 2.1), created the other federal courts (so-called inferior courts), and set their jurisdiction (both original and appellate). The first Congress also made good on the Federalists' promise to include a bill of rights.

Approximately 125 poorly constructed and often overlapping amendments were submitted for congressional consideration. James Madison and George Mason took the lead in editing, rewriting, combining, and crafting the amendments. On September 25, 1789, two thirds of both houses of Congress accepted the conference committee's draft of proposed constitutional amendments. Those amendments were then submitted to the states for

ratification. On December 15, 1791, 10 of the 12 proposed amendments received the approval of the necessary three fourths of the states (as stipulated in Article V) to become part of the Constitution. These first 10 amendments became known collectively as the Bill of Rights.

Ironically, the first two proposed amendments did not pass. The very first proposed amendment, if ratified, would have called for one member of Congress for every 30,000 people. Had the original first amendment been ratified, there would be more than 10,130 members in the U.S. House of Representatives (based on an estimated population of 304 million). Obviously, that would be an unwieldy and unworkable assemblage. The original second amendment proscribes members of Congress from receiving pay increases until after the subsequent congressional election. Although rejected in 1791, this amendment was subsequently ratified by the states (38 states, or three quarters of the 50 states), thus becoming the 27th and last constitutional amendment on May 18, 1992—201 years after it was first rejected.

Conclusion

The Constitution of the United States is now more than 220 years old. It is the oldest functioning nation-state constitution in the world. This great document did not just happen. It was crafted out of necessity and built on the experiences and philosophies of many interrelated people. The document is quite short; it contains a preamble or short justification, 7 articles, and 27 formal alterations. Altogether, the document is just 7,606 words long, far fewer than the words in most daily newspapers' sports sections. But the Constitution's words mean something. They contain essential legal concepts that guide and direct the way the American people function, preserve their culture, and settle differences. It is the rock on which the American political system is built. Perhaps Henry Clay (1850) said it best: "The Constitution of the United States was made not merely for the generation that then existed, but for posterity—unlimited, undefined, endless, perpetual posterity" (para. 1).

References and Further Readings

- Allen, W. B., & Lloyd, G. (1985). *The essential antifederalist*. Lanham, MD: University Press of America.
- Amar, A. R. (2005). *America's constitution: A biography*. New York: Random House.
- Annenberg Learning Center: <http://www.learner.org/biographyof/America.com>
- Articles of Confederation. (1777). Available at <http://www.usconstitution.net/articles.html>
- Avalon Project: <http://www.yale.edu/lawweb/avalon/avalon.htm>
- Beard, C. A. (1913). *An economic interpretation of the Constitution*. New York: Macmillan.
- Bernstein, R. B., & Agel, J. (1992). *Amending America*. New York: Random House.
- Boutell, H. (1896). *The life of Roger Sherman*. Chicago: A. C. McClurg.
- Bowen, C. D. (1986). *Miracle at Philadelphia: The story of the Constitutional Convention, May to September 1787*. Boston: Little, Brown.
- Brown, R. E. (1956). *Charles A. Beard and the Constitution*. Princeton, NJ: Princeton University Press.
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Cappon, L. J. (Ed.). (1987). *The Adams Jefferson letters*. Chapel Hill: University of North Carolina Press.
- Center for Civic Education: <http://www.civic.org>
- Clay, H. (1850). *Speech, U.S. Senate*. Available at http://www.jrank.org/quotations/pages/329/Henry_Clay.html
- Collier, C., & Collier, J. L. (1986). *Decision in Philadelphia: The Constitutional Convention of 1787*. New York: Ballantine.
- Declaration of Independence. (1776). Available at <http://www.usconstitution.net/declar.html>
- Elkins, S., & McKittrick, E. (1962). *The founding fathers: Young men of the Revolution*. Washington, DC: Service Center for Teachers of History.
- Ellis, J. (2002). *Founding brothers: The revolutionary generation*. New York: Vintage.
- Farrand, M. (1913). *The framing of the Constitution of the United States*. New Haven, CT: Yale University Press.
- Federalist No. 51. (1788). *The federalist papers*. Available at http://www.law.ou.edu/ushistory/federalist/federalist_50_59/federalist.51.shtml
- Franklin, B. (1787). *Benjamin Franklin requests prayer in the Constitutional Convention*. Available at <http://www.constitution.org/primarysources/franklin.html>
- Griswold v. Connecticut, 381 U.S. 479 (1965).
- Hall, K. L., Wiecek, W. M., & Finkelman, P. (1996). *American legal history: Cases and materials* (2nd ed.). New York: Oxford University Press.
- Hamilton, A., Madison, J., & Jay, J. (1961). *The federalist papers*. New York: New American Library.
- Hobbes, T. (1904). *Leviathan* (A. R. Waller, Ed.). Cambridge, UK: Cambridge University Press. (Original work published 1651)
- Hobbes, T. (1909 1914). *Of man, being the first part of Leviathan*. Available at <http://www.bartleby.com/34/5/13.html>
- Hofstadter, R. (1948). *The American political tradition and the men who made it*. New York: Random House.
- Jensen, M. (1940). *The Articles of Confederation: An interpretation of the social constitutional history of the American Revolution, 1744 1781*. Madison: University of Wisconsin Press.
- Ketchman, R. (Ed.). (1986). *The anti federalist papers; and, the Constitutional Convention debates*. New York: New American Library.
- Kurland, P. B., & Learner, R. (Eds.). (1987). *The founders' Constitution*. Chicago: University of Chicago Press.
- Lee, R. H. (1776). *Independence resolution*. Available at <http://www.usconstitution.net/leeindep.html>
- Library of Congress: http://www.americaslibrary.gov/cgi_bin/page.cgi
- Locke, J. (1764). *Two treatises of government*. London: A. Millar. (Original work published 1689)
- Magna Carta. (1215). Available at <http://www.fordham.edu/halsall/source/magnacarta.html>

- Martin, L. (1911). *The origin and growth of the American Constitution*. Boston: Houghton Mifflin.
- McCulloch v. Maryland, 17 U.S. 316 (1819).
- McDonald, F. (1976). *We the people*. Chicago: University of Chicago Press.
- McDonald, F. (1985). *Novus ordo seclorum: The intellectual origins of the Constitution*. Lawrence: University of Kansas Press.
- Missouri v. Holland, 252 U.S. 416 (1920).
- Montesquieu, C., Baron de Secondat. (1989). *The spirit of the laws* (A. M. Cohler, B. C. Miller, & H. S. Stone, Eds. & Trans.). Cambridge, UK: Cambridge University Press. (Original work published 1748)
- Patrick, J. (2005). *The Bill of Rights: A history in documents*. New York: Oxford University Press.
- Peltason, J. W., & Davis, S. (2000). *Corwin & Peltason's Under standing the Constitution*. New York: Harcourt Brace.
- Pennsylvania v. Nelson, 350 U.S. 497 (1956).
- Rakove, J. N. (1979). *The beginnings of national politics: An interpretive history of the Continental Congress*. New York: Alfred Knopf.
- Rakove, J. N. (1997). *Original meanings: Politics and ideas in the making of the Constitution*. New York: Vintage.
- Resolution of Congress. (1787). Available at http://www.consource.org/index.asp?bid_582&documentid_71375
- Roe v. Wade, 410 U.S. 113 (1973).
- Rossiter, C. (1966). *1787: The grand convention*. New York: Macmillan.
- Rousseau, J. J. (1997). "The social contract" and other later political writings (V. Gourevitch, Trans.). Cambridge, UK: Cambridge University Press.
- Storing, H. J. (1981). *What the anti federalists were for*. Chicago: University of Chicago Press.
- Sunstein, C. R. (2001). *Designing democracy: What constitutions do*. Oxford, UK: Oxford University Press.
- Tribe, L., & Dorf, M. C. (1991). *On reading the Constitution*. Cambridge, MA: Harvard University Press.
- University of Virginia: <http://www.virginia.edu/pjm/home.html>
- U.S. Constitution. (1787). Available at <http://www.usconstitution.net/const.html>
- Vile, J. R. (1996). *Encyclopedia of constitutional amendments and amending issues*. Santa Barbara, CA: ABC CLIO.

81

URBAN POLITICS

JUNE SAGER SPEAKMAN

Roger Williams University

Summarizing and synthesizing the literature on urban politics is a challenging and rewarding task. The field is by definition interdisciplinary, encompassing political science, economics, sociology, planning, and other fields. To their credit, political scientists have effectively integrated the work of these allied scholars into their research and writing on urban governance.

This chapter focuses on the study of urban politics in the United States. This field is particularly vibrant in the United States, given the nation's federal structure and its deep tradition of, and affection for, local government. In the United States, cities have a political life of their own, independent of the national government, and thus are susceptible to fruitful academic study and analysis.

The theoretical approach to urban politics is discussed first, followed by a description of methodological approaches used by political scientists in their research and analysis of cities. As with every topic in political science, the study of cities involves multiple subjects: governance, leadership, and management; elections and participation; and public policy, including land use, infrastructure, transportation, housing, law enforcement, education, economic development, and wealth and poverty. Each of these areas is considered in turn.

Theories of Urban Politics

The number of seminal texts that offer theories of urban politics is relatively small. These texts, however, have had

a major impact on the field of political science, in particular on the analysis of domestic politics. Several of the major conceptual frameworks used by scholars to understand the operation of political systems have been developed through the study of American cities. Elitism, pluralism, and hyperpluralism, to name just the most noteworthy, are theories studied worldwide and applied to many different types of systems and organizations. These theories emerged from the study of Atlanta, New Haven, and other cities. From these foundational texts have emerged several variants offered by political scientists, sociologists, and economists that elaborate or refine the initial offerings. Among these are the growth coalition perspective and regime theory.

Floyd Hunter's (1953) study of Atlanta, *Community Power Structure*, first drew readers' attention to the fact that it may not be the mayor and the City Council who hold all or any of the power in a city. Using sociological methods of reputational analysis and extended interviews, Hunter's examination of the power structure in Atlanta led him to conclude that power was concentrated in the hands of very few people, primarily members of the economic elite.

In *Who Governs?* political scientist Robert Dahl (1963) reacted to Hunter's theory with a detailed examination of New Haven. For Dahl, the way to understand what is happening in a city is to examine the decisions made by policymakers and track the influences on those decisions. From this analysis, Dahl concluded that what happens in

city government is the result of competition among groups in different policy arenas. Power is fragmented and decentralized among the groups, and the same groups do not always win. Rather, the outcomes will vary from policy area to policy area. Land-use decisions may be dominated by property owners, while education decisions are dominated by parents and teachers' unions. The mayor is the mediator among the contending groups. Competition for control of public policy among groups within a city, for Dahl and his pluralist followers, defined urban politics.

In the late 1970s, Douglas Yates and others examined New York City, New Haven, San Francisco, and several other cities and developed an extension of Dahl's theory, variously referred to as *street fighting pluralism*, *hyperpluralism*, or *neopluralism*. As a result of the increased activism in American cities spurred by the civil rights and antiwar movements of the 1960s, power in cities was now so "fragmented, unstable and reactive" that cities had become "ungovernable" (Yates, 1980, p. 85).

A variant of elite theory that gained prominence in the 1970s and continues to be influential today is the growth coalition or growth machine perspective. Neo-Marxist thinker Harvey Molotch and his collaborator, John Logan, provide a seminal contribution to this school of thought. According to Molotch's (1976) article "The City as Growth Machine," landowners and businesses work with government to "intensify the economic functions of land use" (p. 313). This collaboration explains "the shape of the city, the distribution of people and the way they live together" (Logan & Molotch, 1987, p. 2). For a more political take on this perspective, Paul Peterson (1981), in *City Limits*, explains city politics and policy as the result of city leaders' desire to retain businesses and population in the face of intercity and global competition. City policy, then, is not about serving the needs of the residents but about maintaining a friendly business climate.

Clarence Stone's (1989) work, *Regime Politics: Governing Atlanta, 1946–1988*, introduces the concept of the urban regime as yet another means for understanding the politics and policies of cities. For regime theorists, city politics is a combination of elitism and pluralism. Stone and his followers allow that power and influence are fragmented but assert that levels of influence are not equal. Elected officials and government institutions do matter, as do voters, but economic actors, making private decisions, have inordinate influence over the urban economy. These political and economic, public and private actors combine in long-term coalitions, or regimes, that determine the shape of policy in the city. These regimes vary in their elements from city to city, thus allowing an explanation of differences among cities.

Jane Jacobs's (1993) seminal work, *The Death and Life of Great American Cities*, while not a political science text, deserves some attention for its influence on thinking about what makes cities work. First published in 1961, the book

stands as a challenge to the urban renewal strategies proposed and implemented in the 1950s and 1960s. These strategies resulted in the creation of the American suburbs and the draining of American cities of many middle-class residents and much of their vitality. Jacobs, writing from the perspective of an urban planner, argues that the physical layout and administrative divisions of a city determine the city's success. She argues for mixed-use buildings, small blocks, and population density. These characteristics, she argues, will foster cities that are economically vibrant and diverse—places that people care about and care for.

From a governance perspective, Jacobs (1993) argues that so-called great cities must reform their structure by dividing themselves into administrative districts, horizontally, as the settlement houses have done, taking responsibility for all services in "a piece of the city." Jacobs also argues that cities are most successful when they have vibrant neighborhoods—small blocks with a mixture of residential and commercial uses. These neighborhoods engage in self-government: They "weave webs of public surveillance to protect strangers as well as themselves; they grow networks of small-scale everyday public life and thus of trust and social control, and they help assimilate children into reasonably responsible and tolerant city life" (p. 156). From these neighborhoods, city dwellers move out into other healthy, vibrant neighborhoods for their jobs, dentists, recreation, friends, shops, and entertainment, and even their children's schools. This is the diversity the city brings to the life of its residents and the diversity that city officials must foster if they wish to govern a safe, thriving city.

In his recent books, *The Rise of the Creative Class* and *Who's Your City?* urban theorist Richard Florida (2002, 2008) argues that it is the characteristics of the people in the city, not its physical layout or its power relations, that determine a city's success and vitality. Florida's research leads him to conclude that new ideas are generated and our productivity increases when we locate close to each other in cities and that cities with an open-minded and tolerant culture are likely to foster creativity and thus breed economic success and a high quality of life (2002, p. xxi). Florida argues that mayors should abandon the traditional tools of economic development—tax breaks for corporate headquarters and massive investments like sports stadiums and convention centers—and instead focus on policies that encourage the "three T's": technology, tolerance, and talent.

Students of politics who seek to understand the operation of cities would be well-advised to follow the advice provided in one of the classic textbooks in the field, Lineberry and Sharkansky's (1978) *Urban Politics and Public Policy*. They recommend a systems approach to the study of urban areas, one that has both macro- and microelements that take into account resources and constraints imposed by the environment, the nature of decision makers,

the array of individual and group actors in the city, and the availability of information and the state of knowledge about policy problems. In short, cities are complex entities that sit in a complex state, national, and global environment. That complexity makes theory building a challenge.

Overall, there is no grand theory, widely accepted by the discipline, of how cities operate. Informed by sociology and economics and guided by the grand concepts of elitism and pluralism, political scientists continue their work on the development of concepts and theory in the field of urban politics.

Research Methodology

Political scientists who study urban politics employ an array of methods in their examination and analysis of cities. There is, however, a strong preference for qualitative methods, in particular case studies and applied policy studies. Quantitative analysis can be found in the field, particularly when data sets are easily available, such as in the areas of criminal justice, economics, and demographics. In his review of the field, DeLeon concludes, “Although sample survey research and secondary analysis of sample survey data continue to dominate mainstream political science methodology, the case study is most favored in urban political research” (cited in Vogel, 1997, p. 20).

The review of the aforementioned theories supports this conclusion. Hunter’s study of Atlanta and Dahl’s study of New Haven paved the way for DeLeon’s study of San Francisco, Stone’s work on Atlanta, Mollenkopf’s work on New York and San Francisco, and Sonenshein’s work on Los Angeles.

Within these cases, a variety of methods are used for gathering data, from the intensive interviews and reputational analysis used by Hunter in Atlanta to the analysis of government documents, content analysis of the media, and the deep and wide empirical observation (sometimes called the soak-and-poke method of research) of Sonenshein in Los Angeles.

The City in the Federal System

American cities do not exist in a political or administrative vacuum. They sit on the geographic territory of a state within the political boundary of the United States of America. These layers of jurisdiction have multiple consequences for the operation of American cities.

The National Government

Although interest in American cities ebbs and flows at the national level, there are hundreds of federal laws and regulations that affect the operation of cities in multiple ways every day. Federal law and policy imposes mandates,

both funded and unfunded, that dictate policy direction and constrain—and strain—fiscal resources at the municipal level. For example, in the area of education, Title I (the Elementary and Secondary Education Act of 1965), now known as the No Child Left Behind Act, provides about \$30 billion in annual funding to states, most of which is passed through to local school districts. These districts must comply with the mandates of the act, including increased student testing, higher student standards, and sanctions for schools that do not meet those standards. Critics argue that the federal dollars are insufficient to cover the costs of the mandates. Cities, and all school districts, however, must shape their spending and their educational policies to meet federal guidelines.

In many other areas, from law enforcement to environmental regulation, the construction of buildings, and community development and housing to health care financing and health care delivery, federal funding provides the national government with a lever to shape city policy. This is desirable from the perspective of uniform national standards and the guarantee of a minimum level of funding nationwide. This relationship, often called fiscal federalism, is criticized by those who favor local control and regional diversity.

A recent example from Maine reveals the difficulty of the national–local relationship. In that state, in late 2009, the North Berwick school board voted to forgo an \$18,000 federal grant for a family planning program at the high school because the federal government requires, as a condition of the grant, that the school health service offer a so-called morning-after pill in circumstances where it is deemed appropriate. Rather than comply with the federal mandate, the school board chose to do without the federal funds (Claffey, 2010).

Other examples of national standards constraining the behavior of municipal governments come from the Supreme Court. The famous *Brown v. Board of Education* case in 1954, for example, prohibited racial segregation in school districting, thus changing the behavior of hundreds of local school boards. The City of Boston, as another example, found the transportation and districting aspects of its school system taken over in 1974 by Federal Judge Arthur Garrity after his finding that the school board of that city was gerrymandering school districts to ensure racial segregation. The Supreme Court has issued dozens of opinions defining what a city may or may not do regarding the display of religious symbols on city property, the granting of religious holidays to city workers, and the integration of religious ceremonies into city celebrations. The Court has also dictated the conduct of city elections, the parameters of a city’s power of eminent domain, and the propriety of city handgun ordinances.

American cities cannot be fully understood without taking into account the policy and funding preferences and the fiscal and judicial levers of control in the hands of national policymakers.

State Governments

Although local governments predate state governments in U.S. history, the creation of the nation under the Constitution ended the independent existence of local governments. In fact, the U.S. Constitution makes no mention of any level of government below the state level, thus setting in place a system of state dominance over local governments that has been enshrined in constitutional law and political practice. In 1868, in the most famous judicial statement of this relationship, Iowa's Chief Justice, John F. Dillon, promulgated the oft-cited Dillon's Rule that asserts that states exert total control over cities and that cities are "mere tenants at will of the [state] legislature" (cited in Judd & Swanstrom, 2006, p. 38). This relationship was cemented by malapportionment in state legislatures that reflected the view that "the average citizen in the rural district is superior in intelligence, superior in morality and superior in self-government to the average citizen of the great cities" (p. 39). As a consequence of this attitude, by 1900, "every state had ensured that no matter how large the cities became, representatives from rural districts would continue to hold a controlling majority in state legislatures" (p. 39). It was not until the Supreme Court's "one man, one vote" decisions of the 1960s that this malapportionment was addressed and urban areas found their representation in state legislatures and the U.S. Congress increasing. By that time, however, Judd and Swanstrom note, it was too late:

The nation's population and therefore the balance of power in national politics were shifting to the suburbs and to the Sunbelt. If cities had gained equal representation in the state legislatures and in Congress decades earlier, the influence of urban voters would have virtually guaranteed . . . funding for mass transit, public housing, urban revitalization, and community health programs. (p. 40)

Instead, cities continue to struggle for state and local attention to these priorities.

Home Rule

Still within this framework of federal mandates, state dominance, and bias toward rural areas, cities do have significant control over their own politics and policies. Starting with Missouri in 1875, states began to integrate into their constitutions provisions for home rule that permitted municipalities to draft their own charters, make their own laws, and raise their own taxes. Today, 44 states have some provision for home rule in their constitutions. It should be noted, however, that although cities do have varied degrees of control over their own governing structures and public policy decisions, urban reliance on state aid and the ongoing ability of states to exercise statutory and regulatory control over cities make governors and state legislators important influences in the life and well-being of American cities.

Urban Political Structures

Understanding urban politics requires taking into account not only federal and state actors but also the formal locations of power and organization of political power in cities themselves.

A Word on the Urban Political Machine

In most American cities from the time of the founding through the end of the 19th century, the form of government mimicked the state and national model of an elected executive: the mayor, with considerable independent power, who governed with an elected district-based legislature, the city council. Political parties were alive and well in American cities and were in fact the key to understanding the dynamics of urban politics. The urban political system in most major 19th-century cities was organized by tightly run political party machines. These machines organized the recruitment of candidates and ensured their election by financial support and voter mobilization. They mobilized voters by providing them with material rewards—jobs, hods of coal, Christmas turkeys—in return for electoral support. And the organizers of the machine, often not elected officials themselves, enriched themselves through government contracts, payoffs, and other corrupt practices. Although the machines are maligned by some historians, political scientists, and journalists, others give them credit for organizing city politics, bringing new voters into the system, and providing for the material needs of the urban poor. Political scientists and political journalists have written dozens of books on the machines of Kansas City (Dorsett, 1980), New York's Tammany Hall (Riordan & Quinn, 1995), Chicago (Rakove, 1976), San Francisco (Bean, 1968), Jersey City (McKean, 1967), and others that give the reader deep insights into the dynamics, advantages, and disadvantages of this significant urban political form.

The Reform Movement

At the beginning of the 20th century, the corrupt practices and amoral politics of machine politicians generated reaction among good government reformers and progressives. They described the machine as follows:

affront to their ethical values . . . violated their conceptions of policymaking by professional management . . . capitalized on social cleavages within the electorate and was therefore also at variance with the reformers' belief in the interest of the community as a whole. (Lineberry & Sharkansky, 1978, p. 124)

Led by muckraking journalist Lincoln Steffens, the growing newspaper and magazine industry began to investigate and uncover the corruption of the urban political bosses. In this climate, in 1894, a group of good government reformers (often called *goo goos*) gathered to found the National Municipal League, which generated dozens of proposals

for altering the structure of urban politics to undercut the power of the machines. Much of what we see of political structure and process in American cities today is the result of the reformers' work.

Strong Mayor Council Form

The oldest form of city government, and—as noted previously—the one most susceptible to domination by a machine, is the strong mayor-city council form. This form mimics the national government of the United States, with a mayor and a legislature, elected separately, with independent but overlapping powers. There are variations on this theme in terms of how many other members of the executive branch are elected (city solicitors, tax collectors, and school superintendents are sometimes elected separately), whether the mayor can veto acts of the council, whether council members are elected at large or by district, the length of terms, and the presence of term limits, among others. The larger and more demographically and economically diverse the city is, the more likely it is to have this model of government. Older cities in the Northeast and Midwest also favor this form. Of the 25 largest cities in the United States, 17 have the strong mayor model. They are New York, Los Angeles, Chicago, Houston, Philadelphia, San Diego (since 2006), Detroit, Indianapolis, Jacksonville, San Francisco, Columbus, Memphis, Baltimore, Milwaukee, Seattle, Boston, and Denver. The 8 that do not are all in the West or South: Phoenix, San Antonio, Dallas, San Jose, Austin, Fort Worth, Charlotte, and El Paso (Strong Mayor-Council Institute, n.d.).

The advantages of the strong mayor model are as follows: There is the concentrated power and control that cities may need in times of crisis; there is a central figure who can mobilize public sentiment; and the mayor can, if successful, mediate the various contending social, ethnic, and economic interests in the city. The disadvantages of this model are that the chief executive is a politician, who may hold the particular interests of his supporters as more important than the interests of the city as a whole. A strong mayor may be more prone to corruption, given that he has greater access to city resources. And he may lack the budgeting, policy, and technical skills needed to manage a complex organization. Despite these problems, however, the Strong Mayor-Council Institute (n.d.) reports that since 1990, nine major cities (Tulsa, St. Petersburg, Fresno, Oakland, Sioux Falls, Spokane, Hartford, Richmond, and San Diego) have adopted this model, while only three (El Paso, Topeka, and Cedar Rapids) have moved in the other direction.

Council Manager Form

That other direction is the council-manager form of municipal government. In this model, the city council is relatively weak and certainly part-time. There usually is no mayor in this model, although often the council president

bears the title of mayor and serves as the ceremonial or symbolic head of the city. The council appoints a full-time city manager who acts as the chief executive officer of the city. The manager writes the budget, hires and fires all department heads, and consults with the council on only those relatively few issues mandated by the city charter (such as adoption of the budget, changes in the tax rates, or the passage of new city statutes). As noted, this model was developed by reformers as an antidote to the perceived and real corruption of the urban political machine. The stated purpose of the model is to put a professional administrator in charge of the city, separating administration from politics, as a way of minimizing corruption, maximizing efficiency and effectiveness of service delivery, and saving tax dollars. The disadvantage of this model is that the most powerful person in the city is not accountable to the public and, as noted previously, not charged with the political tasks of mobilizing public support or mediating conflict among contending city interests.

The council-manager form of government is seen by its proponents as providing the best of both worlds: a council that is responsive to the public and can hire and fire the manager at will and a manager who is a professional urban administrator and is politically neutral as he carries out the business of the city. For these reasons, according to the International City Managers Association (ICMA, the leading proponent of this model), 3,302 cities have this type of government, as do almost 400 counties, and on average, 63 communities (all relatively small in size) adopt this model each year, while only two abandon it.

Mayor Council Chief Administrative Officer

In recent years, some of the largest American cities, including New York, Los Angeles, Philadelphia, San Francisco, and New Orleans, have seen the addition of a chief administrative officer or a city administrator to the mayor's staff. The position is similar to a city manager in that he or she is hired for his or her professional competence in the administrative aspects of urban management. This model is in some ways a hybrid between the strong mayor model and the council-manager model in that it injects a trained manager into city government while retaining the political functions of the mayor. Given the increased complexity and interdependence of city services and the large number of federal and state policy mandates that bring budgetary and reporting complexity with them, one might expect to see more mayors adding this position to their teams (Fahim, 2005).

Commission Form

The commission form of government was first adopted in Galveston, Texas, in 1900 as a way of increasing the governing capacity of the city council (in the wake of a devastating hurricane) by giving each council member (now called a commissioner) administrative responsibilities.

Five to seven commissioners are elected by the voters. Each commissioner has both executive and legislative responsibility for several policy areas within the city. There is no overall integrative mechanism. For this reason primarily, very few cities employ this model today. Portland, Oregon, is the only large city in this category and is joined by several dozen smaller communities, mostly in the South.

Urban Elections

The timing and structure of elections at the local level are even more varied than the forms of government. As with the move to council-manager government, progressive reformers of the early 1900s advocated successfully for a number of reforms in the electoral process that remain in effect today. The purpose of most of these reforms was to squeeze out of local elections as much politics as possible. Many communities adopted off-year elections for their cities in order to separate the fate of the city from the state and national contests. Nonpartisan elections were widely adopted as well, in keeping with the axiom that “there is no Democratic or Republican way to sweep the streets.” Today, 75% of municipal elections are nonpartisan. A related reform was the implementation of at-large, or citywide elections to replace district or ward elections. This, too, was designed to break the hold of the local political party machine on city politics and to allow for the election of the most qualified candidate regardless of his or her partisan affiliation or address. Other reforms include the adoption of initiative, referendum, and recall as mechanisms for voters to exert direct control over policy outcomes and political leaders in their cities.

Voter Turnout

Although these reforms did have the effect of disconnecting urban elections from national and state political trends, they also resulted in dramatic reductions in voter participation in urban elections. It is not unusual to find turnouts of 10% or lower when only local offices are up for grabs. For example, in the June 2009 elections for five municipalities in Clark County, Nevada, local officials predicted a turnout rate of 12% at best (Hansen, Twitchell, & Van Oot, 2009).

Nonpartisan Elections

Nonpartisan elections also have an impact on both voter choice and voter turnout. Without political parties to organize their thinking and provide cues about candidates, voters may be more likely to stay home and are certainly more likely to be susceptible to other signals when determining their choices at the polls. Raymond (1992) has found that in nonpartisan elections, voters tend to focus on “candidates’ personal qualities, background characteristics, name

recognition and local activism—and not on issues” (p. 248). Ethnic voting increases in the nonpartisan context as voters with no other information use the candidate’s last name to guide their choice. Class bias may also be deepened in nonpartisan elections, given that media advertising becomes that much more important without party labels and party workers. Wealthier candidates thus gain greater access to voters and have an advantage on Election Day. New York City mayor Michael Bloomberg’s 2003 proposal to switch New York City to a nonpartisan system was opposed and ultimately defeated by a coalition of blacks, Latinos, and Asian Americans who argued that such a system would disadvantage candidates and voters from their communities.

At Large Versus District Elections

At-large elections have the advantage of reducing the parochial, neighborhood focus of ward or district representatives. Such a system makes it easier for city council members to consider policy from a citywide perspective and to engage in integrative thinking as they approach the city’s policy challenges. Another argument for this feature is that a broader candidate base (citywide) allows for higher quality candidates. In the United States, 64% of cities, primarily the smaller and more affluent ones, use this model. Opponents of an at-large system argue that it allows for the underrepresentation of ethnic and racial minorities since they are consistently outvoted by majorities. Further, its opposite—district elections—produces politicians more in touch with local concerns. On the other hand, district elections may increase conflict on the council and make compromise on policy difficult. District elections for council are in place in 14% of American cities—mostly larger, diverse cities. Some cities (21%) have adopted a hybrid model in which several members of the council are elected from districts and several at large.

A 2009 lawsuit in Irving, Texas, underscores the significance of this structural feature of urban elections. Of the nearly 200,000 residents of this city, 43% are Hispanic, yet all members of the town council are white. The council is elected on an at-large basis. Resident Manuel Benavidez filed a voting rights suit in federal court arguing that the white voters have used the system to effectively block the election of Hispanic voters. A judge agreed and ordered Irving to adopt a hybrid system in which five council members are elected by district and three are elected at large (Formby, 2010a). The first election under this new system took place on May 8th, 2010. It was expected to produce a reallocation of political power in this Texas city as at least one district was gerrymandered to include a majority of Hispanic voters. Ironically, voters in that district elected an African American, while voters in one of the remaining at large districts elected the first-ever Hispanic council member, who defeated a long-term white incumbent (Formby, 2010b).

Initiative, Referendum, and Recall

City charters may have provisions for the direct-democracy instruments of initiative, referendum, and recall. The first is a voter-initiated ballot measure (shall the class size in city schools be capped?), the second is a council-initiated measure (shall the city permit casino gambling or not?), and the third is a voter recall of an elected official. In 1997, for example, the voters of Minneapolis voted to limit city financial assistance to any new sports facility. In the mid-1980s, the voters of San Diego enacted a number of antigrowth measures. In many cities, the city charter requires that sales of city property over a certain value, or increases in taxes, be placed before the voters. Finally, in 2009 alone, recall petitions were filed against the mayors of Akron, Flint, Kansas City, and Portland and against several council members from San Jacinto and San Diego.

The advantages of these measures from a democratic theory perspective are clear. Voters get to make choices directly instead of through their elected representatives, thus increasing voter contact with public policy and enhancing the legitimacy of the outcome. The disadvantages are that recalls may be used by the losers in the most recent election to have a second bite at the apple; initiative and referendum campaigns can be dominated by well-funded interest groups (as in most casino measures) or, alternatively, result in tyranny of the majority (as in some anti-immigrant or anti-gay-marriage measures). And again, given the propensity for low interest in and low turnout for local elections, the ability of small numbers of voters to make major policy decisions is a significant possibility.

Clearly, there is limitless variation in the political structures and processes of American cities. No one form is accepted by political scientists as the ideal form for a city. What political scientists can offer is an understanding of the consequences of the adoption or change in a particular form or process for efficiency of service delivery and representative democracy.

Race and Ethnicity in American Cities

Urban areas are always the most diverse parts of any country since urban ports provide the destination for immigrants and the urban economy offers opportunities to job seekers from other parts of the country and the world. Between 1820 and 1919, 33.5 million immigrants moved to the United States; of those, at least 75% stayed in cities. Internal migration patterns are equally dramatic: Between 1910 and 1970, more than 6 million blacks left Southern fields for Northern cities (Judd & Swanstrom, 2006).

After World War II, the development of American suburbs was fostered by the GI Bill's provision for easy mortgages

for war veterans, the construction of the interstate highway system, the proliferation of the automobile, and the gradual replacement of urban factory jobs with beltway service and technology jobs. These changes, along with the urban riots of the 1960s and Supreme Court decisions prohibiting school segregation and, in some cases, mandating busing to integrate schools, led to the phenomenon known as *white flight*, the dramatic migration of middle-class whites to the suburbs.

The consequence of these demographic shifts is a large number of majority-minority cities. According to the 2000 Census, in several major cities, including Detroit, Birmingham, New Orleans, Baltimore, Memphis, and Washington, D.C., the African American population exceeds 50%. In still others, including San Antonio, El Paso, Brownsville, Corpus Christi, Laredo, Hialeah (Florida), Miami, Paterson (New Jersey), Reading (Pennsylvania), Downey (California), and East Los Angeles, the Hispanic population comprises more than half of the city.

Unfortunately, the racial segregation of America's urban areas has been accompanied by economic segregation and all of the social consequences of urban poverty.

Wealth and Poverty in American Cities

In the United States, unlike most other countries, household incomes rise as one moves away from the city center. (Paris stands as the most notable model of the opposite dynamic: The wealthy live at the city's center, and the poor live in the suburbs that ring the city.) Dreier, Mollenkopf, and Swanstrom (2001) present data that show that the gap between per capita income in 85 cities and their surrounding suburbs grew continuously between 1960 and 2000. This gap is the result of the migration of two-parent households and new service-sector jobs to the suburbs, while single-parent, female-headed households and job loss stayed behind in the cities (Dreier et al.).

As the demographic shifts occurred throughout the 20th century and racial and economic segregation intensified, conditions in inner cities deteriorated to a point that, by 1967, the National Commission on Civil Disorders (also called the Kerner Commission) declared the existence of "two nations—one black, one white—separate and unequal" (Judd & Swanstrom, 2006, p. 244). A number of books and articles written in the 1970s, 1980s, and 1990s revealed the existence of an "urban underclass" or a "new urban poverty" driven by persistent unemployment and related to family disorganization, drug use, teen pregnancy, high high school drop-out rates, neighborhood deterioration, and high crime. Jargowsky found that between 1970 and 1990, the "spacial concentration of the poor rose dramatically in many U.S. metropolitan areas," as the poor became "more physically isolated from the social and economic mainstream of society" (cited in Berube, Katz, & Lang, 2005, Vol. 2, p. 138).

According to the 2000 census, nationwide, urban dwellers are more than twice as likely to be poor as are their suburban counterparts (Berube et al., 2005). Some of the highest urban poverty rates were in Hartford (30.6%), Miami (28.5%), and Newark (28.4%) (Berube et al.). The good news of the 2000 census was that the booming economy of the 1990s reversed the decades-long trend of the concentration of poverty in the inner cities of America. Whether that good news is sustained through the first decade of the 21st century remains to be seen.

Urban Public Policy

Discussions of public policy in the urban context occur on two levels: on the ground, where mayors and their administrators deliver services to city dwellers, from education to crime-fighting to trash removal to economic development; and from above, as politicians in Washington, D.C., develop (or not) a national approach to the nation's cities.

On the Ground

The academic analysis of policy development and delivery at the level of individual cities takes place primarily in the subfield of public administration, with its focus on city management, public finance, economic development, and urban planning. Here, the discussions and analyses are found in the pages of such publications as *Public Administration Review* and *Governing* magazine; in the various forums provided by the ICMA; or at annual meetings of city treasurers, tax collectors, personnel directors, or urban planners. Often, these discussions are among both scholars and practitioners and are focused on debates about the techniques of governance. Any student of urban politics would be well served to listen in on these conversations. Sound management, attentiveness to best practice, familiarity with new technology, and the use of analytical tools like benchmarking and strategic planning—these can often make the difference between a city that works and one that doesn't. Consider, for example, the introduction of CompStat into the operation of the New York City police department. This data-driven approach to tracking and preventing crime has been credited with contributing to a dramatic reduction in the city's murder rate, from 1,181 in 1995 to 596 in 2003. By 2004, as a result of New York's success with this tool, at least one third of the nation's largest police departments had adopted a similar approach (Weisburd, Mastrofski, Greenspan, & Willis, 2004).

Again, the work in this field is primarily practical, aimed at discovering what works and what doesn't, and at connecting policymakers with each other to facilitate the improved management of American cities and towns.

Think tanks like the Brookings Institution and the Urban Institute provide a layer of academic sophistication to the analysis of the ground-level operation of American

urban centers. For example, in 2008, Brookings released a study of Chattanooga that examined how sound planning and private-sector partnerships combined to rekindle economic growth in this small city. Here again, students of cities are well served by this literature, which focuses on the factors that help illuminate explanations for success and failure in urban policy.

In Washington

Urban policy at the national level is much more variable than at the local level. Presidents and Congress focus on cities when they must (as in the 1960s in the midst of urban riots) or when the political winds blow in that direction. Urban policy is essentially the province of the Democratic Party, which captured urban voters as part of the New Deal coalition and has yet to let them go. The 1960s was the decade when attention to urban issues reached a peak in Washington, pushed by events on the ground and pulled by Democratic politicians like Lyndon Johnson, whose War on Poverty promised to "eliminate the paradox of poverty in the midst of plenty in this nation," and whose Model Cities legislation claimed that "improving the quality of urban life is the most critical domestic problem facing the United States" (Judd & Swannstrom, 2006, p. 178). Throughout the 1970s, support for these programs waned until, by the Reagan presidency, the national debate focused on eliminating the programs. Even President Clinton, a Democrat, rarely spoke of urban policy and programs, but rather focused on welfare reform, immigration reform, and fiscal responsibility. State and federal aid to cities reflects this trend: In 1978, 26% of municipal revenues came from state and federal sources; by 2000, the share had dropped to 7% (Judd & Swannstrom). Cities were on their own. The election of Barack Obama may signal renewed attention to America's cities with the establishment of the Office of Urban Affairs as one of his first acts as president (White House, 2009).

Conclusion

American cities are fruitful ground for political scientists. They have all of the political phenomena of the national level, from the institutions of power to the processes of democracy to the outputs of government. Their relatively small scale permits in-depth analysis, and their infinite variety provides a wealth of data for comparative analysis. Plentiful resources exist to assist the political science researcher, from detailed census data to endless government documents to journalist case studies and professional journals. Theories abound as to what makes cities tick. If President Obama is correct that "the economic health and social vitality of our urban communities are critically important to the prosperity and quality of life for Americans" (White House, 2009, Section 1), then political

scientists have an important role to play in analyzing how cities work well and what can be done to make them work better.

References and Further Readings

- Bean, W. (1968). *Boss Reuf's San Francisco*. Berkeley: University of California Press.
- Berube, A., Katz, B., & Lang, R. E. (Eds.). (2005). *Redefining urban and suburban America* (Vols. 1-3). Washington, DC: Brookings Institution.
- Claffey, J. (2010, January 8). "Morning after" pill exclusion blocks federal funds for Noble High School. *Fosters.com*. Retrieved January 10, 2010, from http://www.fosters.com/apps/pbcs.dll/article?AID=/20100108/GJNEWS_01/701089874
- Dahl, R. (1963). *Who governs?* New Haven, CT: Yale University Press.
- Dorsett, L. W. (1980). *The Pendergast machine*. Omaha: University of Nebraska Press.
- Dreier, P., Mollenkopf, J., & Swanstrom, T. (2001). *Place matters*. Lawrence: University Press of Kansas.
- Eichenthal, D., & Windenecht, T. (2008, September). A restoring prosperity case study: Chattanooga, Tennessee. *Brookings*. Retrieved December 20, 2009, from http://www.brookings.edu/papers/2008/0917_chattanooga_eichenthal_windeknecht.aspx
- Fahim, M. (2005, December 18). Council managers are running more and more American cities. *City Mayors Government*. Retrieved December 20, 2009, from http://citymayors.com/government/council_managers.html
- Florida, R. (2002). *The rise of the creative class*. New York: Basic Books.
- Florida, R. (2008). *Who's your city?* New York: Basic Books.
- Formby, B. (2010a, January 8). Irving moves closer to single member city council districts. *Dallas News*. Retrieved January 10, 2010, from <http://www.dallasnews.com/sharedcontent/dws/news/city/irving/stories/010910dnmetirvingdistricts.5ac2b662.html>
- Formby, B. (2010b, May 16). *May 8 election offers glimpse of Irving's new system at work*. Retrieved May 20, 2010, from http://www.txn.com/sharedcontent/dws/news/localnews/stories/DN_irvanalysis_16met.ART0.Central.Edition1.a1f95d4.html
- Governing: <http://governing.com>
- Hansell, W. (2001, Spring). Council manager government: Alive and leading today's best managed communities. *National Civic Review*, 90, 41-43. Retrieved December 20, 2009, from http://www.ncl.org/publications/ncr/90_1/Hansell.pdf
- Hansen, K., Twitchell, J., & Van Oot, T. (2009, June 2). Polls close amid light municipal turnout. *Las Vegas Sun*. Retrieved December 20, 2009, from http://www.lasvegassun.com/news/2009/jun/02/voters_head_polls_select_mayors_council_members_ju/
- Hunter, F. (1953). *Community power structure*. Chapel Hill: University of North Carolina Press.
- International City Managers Association: <http://icma.org>
- Jacobs, J. (1993). *The death and life of great American cities*. New York: Random House.
- Judd, D., & Swanstrom, T. (2006). *City politics* (5th ed.). New York: Pearson Education.
- Lineberry, R., & Sharkansky, I. (1978). *Urban politics and public policy* (3rd ed.). New York: Harper & Row.
- Logan, J., & Molotch, H. (1987). *Urban fortunes*. Berkeley: University of California Press.
- McKean, D. (1967). *The boss*. New York: Russell & Russell.
- Mollenkopf, J. (1983). *The contested city*. Princeton, NJ: Princeton University Press.
- Molotch, H. (1976). The city as growth machine. *American Journal of Sociology*, 82, 309-332.
- Peterson, P. (1981). *City limits*. Chicago: University of Chicago Press.
- Rakove, M. (1976). *Don't make no waves . . . don't back no losers*. Bloomington: Indiana University Press.
- Raymond, P. (1992, April). The American voter in a nonpartisan, urban election. *American Politics Quarterly*, 20, 247-260.
- Riordan, W., & Quinn, P. (1995). *Plunkitt of Tammany Hall*. New York: Signet.
- Sonenshein, R. (2004). *The city at stake: Secession, reform and the battle for Los Angeles*. Princeton, NJ: Princeton University Press.
- Stone, C. (1989). *Regime politics*. Lawrence: University of Kansas Press.
- Strong Mayor Council Institute. (n.d.). *Structural trends: U.S. cities*. Retrieved November 15, 2009, from <http://strongmayorcouncil.org/structuraltrends.html>
- Vogel, R. K. (Ed.). (1997). *Handbook of research on urban politics and policy in the United States*. Westwood, CT: Greenwood.
- Weisburd, D., Mastrofski, S. D., Greenspan, R., & Willis, J. (2004). The growth of CompStat in American policing. *Police Foundation Reports*. Retrieved January 8, 2010, from <http://www.policefoundation.org>
- White House. (2009, February 9). *Executive order creating the Office of Urban Affairs*. Retrieved December 20, 2009, from http://www.whitehouse.gov/the_press_office/Executive_Order_Establishment_of_the_White_House_Office_of_Urban_Affairs
- Yates, D. (1980). *The ungovernable city*. Cambridge: MIT Press.

82

MEDIA AND POLITICS

BAS W. VAN DOORN

The College of Wooster

The question of how the media affect politics is complicated, but in its most basic formulation, the correct answer to it is this: It depends. Although early conventional wisdom held that the media had strong, direct, so-called hypodermic effects, more recent research provides convincing evidence that individual-level and contextual factors significantly influence the extent to which media affect people's political behavior and beliefs and, eventually, public policy. The intent of this chapter is to serve as an introduction to some of the basic theories, insights, and debates about media and politics. In doing so, it touches on issues of media ownership, media bias, politicians' use of media, scholarly models of media effects, and new media. The final section of the chapter discusses some fruitful areas for future research.

Media Ownership, Consolidation, and the Marketplace of Ideas

The Role of Media in Democracies

The media are commonly understood to be able to—and many would argue, obligated to—provide a forum for the expression and discussion of a diverse range of oftentimes conflicting ideas. This is especially important in the run-up to elections, where citizens are to base their decisions at least partially on whose policy proposals they

deem most attractive. However, even when elections are far off, this type of discussion serves to inform citizens, and most scholars believe that democracy benefits from a more knowledgeable citizenry.

This forum function of the media is often defended from the perspective of truth seeking, the argument being that only through the meeting of diverse opinions will the truth emerge. This is known as the *marketplace of ideas* perspective. There are at least two main lines of criticism to this approach. First, many would contest the notion of truth seeking in politics. The meat of politics, this perspective suggests, is value conflict, which in the final analysis is irreconcilable. Thus, there is no underlying truth or fundamental consensus that may be exposed through discussion.

Nonetheless, it seems uncontroversial that most debates benefit from the provision of factual information and from discussion of what implications different value priorities have in the policy area at hand. In modern societies, this would be impossible without the media. There are, however, profound concerns about the extent to which the media provide the information and perspective necessary to an intelligent consideration of the merits and implications of policy. This lies at the root of an objection to the marketplace of ideas as applied to modern societies: Even if an uncontroversial truth did exist, today's media coverage of the issues is not conducive to helping citizens divine it. In Graber's (2003) words, "The cacophony of

voices in today's marketplace of ideas often confuses non-experts more than it enlightens them" (p. 144). In the electoral context, the media are also widely criticized for a lack of attention to substance, in favor of the horse race and strategic dimensions of electoral politics (Patterson, 1993).

Ownership and Regulation

This is not to say that democracy would be better off without the media or without a marketplace of ideas. It merely indicates that the current mass media are not serving the public very well. One proposed reason for this, especially in the U.S. context, pertains to media ownership. What sets U.S. media apart from other media systems is private ownership and the relative independence from government regulation.

Worldwide, there is much variation in terms of public versus private ownership, but one can distinguish among three basic models of media ownership: purely public, mixed, and purely commercial (Iyengar & McGrady, 2007). In general, public broadcasting is publicly funded, through either fees or taxes. In return for public funding, public broadcasters are required to provide a certain level of public affairs programming. With the emergence of cable, very few countries still have purely public television markets, having transitioned instead to a mixed model featuring both public and commercial stations.

Public ownership is associated with some clear patterns. Content analysis shows that levels of political content are significantly higher on public channels than on commercial channels (Krüger, 1996). Moreover, comparisons of politics-related coverage on public and commercial outlets indicate that the former provide more substantive coverage and higher proportions of internationally oriented news (Heinderyckx, 1993).

There are also some noted potential drawbacks to public ownership. First, some people resent the expenditure of taxes or the levying of fees for public programming. Second, there are concerns about the compromised independence of publicly owned channels that, depending on government funding, may be disinclined to bite the hand that feeds them, thus resulting in an uncritical attitude toward the powers that be. Although this is most definitely the case in undemocratic regimes, there is no systematic evidence that press freedom is compromised by public ownership in Western democracies. In fact, some studies suggest that press freedom and public ownership are positively correlated in democracies (Iyengar & McGrady, 2007).

Media Consolidation

One of the most salient developments in the media landscape is consolidation of ownership into media conglomerates. Bagdikian (1997) demonstrates that fewer

and fewer cities have competing daily newspapers: In 1920, 700 cities had such competition, a number that had dropped to 19 by 1996. Independent ownership of newspapers has also declined precipitously, dropping from 83% in 1940 to 24% in 1990. A high and increasing proportion of newspapers is now owned by a limited number of newspaper chains such as Knight-Ridder. This raises concerns about the extent to which there is a viable marketplace of ideas in one-paper towns, both because a larger number of outlets will likely reflect a wider range of ideas and because a lack of competition plausibly reduces incentives for quality reporting. Television networks have likewise been acquired by large companies such as the Disney Corporation (ABC), General Electric (NBC), Viacom (CBS), NewsCorp (Fox), and Time-Warner (CNN).

Whereas network television executives used to expect and accept the lack of profitability of news shows, counting on it to be offset by more profitable entertainment programming, the networks' new owners have no such perspective, expecting all programming to be as profitable as possible. This has profoundly, and many would argue negatively, affected the production process and content of network news. First, to reduce costs, staff has been cut and foreign bureaus have been closed. Second, the content of the news has shifted from traditional so-called hard news to a blend of entertainment and news known as soft news. The primary goal of news is now to entertain rather than inform. A third hypothesized effect of consolidation is simultaneous censorship of news that reflects badly on the parent company and encouragement of news that reflects well on the company (Erikson & Tedin, 2007).

Media Bias

Irrespective of how the news is reported, the media are accused of being biased and lacking objectivity. Left-oriented media critics argue that as businesses, especially in the era of consolidation, the media are dependent on advertising revenue and are thus inclined to tilt to the right and support the status quo. Critics on the right, on the other hand, point to the disproportionate number of liberals and Democrats among journalists as evidence for left-oriented bias in the news. It is impossible to argue with the numbers on this point: Taken as a group, journalists are indeed significantly more liberal and more likely to identify as Democrats. The question remains, however, whether these ideological and partisan preferences translate into biased reporting. The answer to this question is not entirely clear. One study reports that a majority of a sample of journalists suspected that their political opinions sometimes affected their reporting (Dautrich & Deneen, 1996). Another experimental study found only a minor effect of political ideology on how

reporters described reporting a hypothetical story (Patterson & Donsbach, 1996).

There are several reasons why the ideological preferences of journalists do not strongly and consistently affect media content. First, journalists take seriously their professional code of ethics. Second, editors, who make the final decisions as to what is published or aired, are responsible to CEOs and other higher-ups in the corporate structure and thus have incentives to make the final media product palatable to these people, who are likely to be more conservative than the reporters.

Important to debates about media bias is the hostile media phenomenon: Irrespective of media content, people see the media as predisposed against their own position and biased toward the opposing camp (Vallone, Ross, & Lepper, 1985). Thus, it may be impossible for the media to be universally regarded as unbiased. Evidence of this is provided in a piece by Dalton, Beck, and Huckfeldt (1998), who find that there is little correspondence between newspapers' ideological leanings and people's perceptions of those leanings.

Politicians' Use of Media

In the age of mass media, politics has changed from a largely interpersonal to a predominantly mediated activity. The public gets virtually all of its political information through the media, and politicians and political groups have adapted to this situation. Media strategies are now part and parcel of electoral campaigning as well as the policy-making process. The most prominent research in this field concerns presidential behavior and campaign advertising, especially the negative variant.

Political Advertising

Most of the research on political advertising has focused on its negative tone and the consequences thereof. The classic work along these lines is Ansolabehere and Iyengar's (1996) *Going Negative*, which argues that negative advertising causes cynicism about politics and, consequently, depresses voter turnout, especially among independents. This has come to be known as the demobilization hypothesis.

It did not take long for an alternative hypothesis to arise. Proponents of the stimulation hypothesis (e.g., Finkel & Geer, 1998) argue that rather than demobilizing voters, negative ads stimulate turnout by raising the stakes of the election, making emotional appeals, providing information, and increasing people's motivation to learn about the candidates.

Some serious difficulties for measurement and operationalization characterize the study of the effects of advertising tone on turnout. These difficulties mostly originate in the subjective nature of the concept of negativity. Some

scholars (Finkel & Geer, 1998) code ads themselves and decide whether they themselves think they are negative or positive. Others (Wattenberg & Brians, 1999) code voters' comments on the negative or positive tone of ads they remember.

Another big challenge in the literature is the measurement of ad exposure. Some scholars (Wattenberg & Brians, 1999) simply rely on recollection of ads as a measure of exposure. The method that is perhaps most prevalent is the use of ad archives to get a picture of which ads were aired in which campaigns. These data are subsequently combined with a media exposure variable to arrive at a measure of ad exposure (Finkel & Geer, 1998). Others (Freedman & Goldstein, 1999) combine indicators of people's media use with rather sophisticated measures of what ads were broadcast, where, and when.

In an effort to establish which of the two sides in the debate is right, Lau, Sigelman, Heldman, and Babbitt (1999) offer a meta-analysis of the literature on this topic. They conclude that there is no reliable statistical basis for the claim that negative advertisements are liked less than positive advertisements. Furthermore, their analysis suggests that there is no evidence that negative ads are more effective than positive ads. Finally, they conclude that the literature thus far does not warrant the contention that negative advertising demobilizes the electorate.

There is thus no compelling evidence that negative advertising hurts voter turnout. Moreover, contrary to conventional wisdom, there is little to back up the claim that negative advertising is particularly effective in persuading people to vote for you (or against your opponent). In fact, one could argue that negative advertising is a boon to the democratic process. Lau, Sigelman, and Rovner (2007), for example, find that negative advertising does not win elections but that it does increase people's knowledge about the issues and tends to be more memorable than positive appeals. Agreeing with Patterson's (1993) contention that media coverage of campaigns, in its emphasis on strategy and poll numbers over substance, divides rather than connects candidates and citizens, Franz, Freedman, Goldstein, and Ridout (2007) suggest that political ads are valuable sources of information for voters.

Going Public

Politicians have also adjusted to the media in their policy-making efforts. In terms of presidential behavior, the most salient phenomenon is going public (Kernell, 1992). Going public refers to a president going over Congress's head and appealing to the public directly, oftentimes via a televised address. The goal is to activate or change public opinion, leading to legislators feeling pressured to fall in line with the president's policy proposal. This strategy is potentially risky, especially for unpopular presidents, and tends to be more prevalent in

times of divided government, when the presidency and the legislative branch are controlled by different parties.

Basic Models of Media Effects

Hypodermic Model

Having discussed some of the literature on what determines media content, it is appropriate to turn to the effects that this content has on citizens. One culturally salient account of media effects holds that those who control the media directly, immediately, and strongly affect what citizens know, believe, and do politically. This model, which has little empirical support, is known as the *hypodermic model of media effects*, since it depicts the media as injecting information and opinion into the unresisting public. Its effects, then, are like that of a drug that is introduced into the bloodstream.

An anecdotal example of such an effect is Orson Welles's 1938 radio broadcast of H. G. Wells's *War of the Worlds*, which featured a realistically staged report of alien invasion in New Jersey. Although subsequent mythmaking has exaggerated the scope and intensity of people's responses to the broadcast, it is beyond dispute that significant numbers of people took the report at face value and genuinely believed aliens were taking over New Jersey, with some of them taking to the streets, calling law enforcement, and heading out to the location of the alleged Martian sighting.

Another factor contributing to, and echoing, the belief in the hypodermic model is the development of sophisticated propaganda during and in between World Wars I and II. Walter Lippmann (1922), one of the pioneering theorists of media effects, was part of U.S. propaganda efforts during World War I, and this experience led him to believe that the media, especially when speaking in a unified voice, wielded tremendous power over public opinion.

Minimal Effects Model

Despite the vivid examples of dramatic media influence and the widespread use of propaganda by governments the world over, disagreement does exist among citizens on virtually all political issues. This suggests that government control over the media is not complete, that the media present a diverse array of opinion, and that people do not simply accept media information as gospel. This, then, indicates a need to think differently about media effects.

One school of thought questioning the hypodermic model developed during the time from the 1940s through the 1960s, in a series of studies interested in explaining whom people vote for (e.g., Lazarsfeld, Berelson, & Gaudet, 1948). These studies were designed in the expectation of demonstrating strong media effects. However, the

surprising finding was that the media had surprisingly little influence. What developed as a consequence is the minimal effects model.

The minimal effects model posits three explanations for the lack of a strong influence of the media on political beliefs and behavior. First, partisanship limits the potential of media effects. Many people identify with one of the two major political parties, causing them to be loyal to this party, irrespective of what the media have to say about it.

Second, and related, people tend to screen incoming information so as to only pay attention to that which comports with their partisan and ideological predispositions. People sought out congenial information and attempted to avoid information coming from the opposing party's campaign. If unable to do the latter, people often misperceived the opposition's statements, so as to make them consistent with expectations.

Third, minimal effects studies documented the importance of interpersonal conversations. Even if media content was potentially persuasive and caused people to vacillate from their original opinions or develop opinions inconsistent with the party line, conversations with fellow partisans often served to make them return to their preexisting opinions or candidate preferences or to change their developing opinion on a new issue to the party line. This has the added effect of strengthening people's partisan allegiances, making future departures from party orthodoxy even less likely.

Subtle Effects: Agenda Setting, Priming, and Framing

Although few scholars dispute the limiting role of partisan identification on the influence of the media, a number of compelling streams of research have provided intriguing evidence contradicting the notion of minimal effects. In its stead, they encourage a view of politically consequential media influence, without heralding back to the simplistic notion of hypodermic effects.

The basic point of view underlying these streams of research is that although the media clearly do not—or only very rarely do—dramatically, uniformly, and instantaneously alter the public's views, there are more modest, yet important, ways in which the media affect the public mind. Three processes have received particular attention: agenda setting, priming, and framing.

Agenda Setting

The concept of agenda setting finds perhaps its most famous and concise expression in Cohen's (1963) claim that the media "may not be successful much of the time in telling people *what to think* [*italics added*], but it is stunningly successful in telling its readers what to think about" (p. 13). In other words, the agenda-setting hypothesis claims that while media content does not have the hypodermic

opinion-changing effect that early research feared it had, it does have great impact in determining what issues people will focus on and judge to be important.

Early work on agenda setting tended to simply compare the issue focus of the mass media in a certain time span with the issues that citizens tended to regard as most important during the same period of time. Although certainly suggestive, these studies do not provide convincing causal evidence of agenda setting. For that reason, many more recent agenda-setting studies have adopted an experimental approach, which allows for a comparison of people exposed to information about certain issues with otherwise identical people who have not had such exposure. Iyengar and Kinder (1987), in *News That Matters: Television and American Public Opinion*, one of the classic works on agenda-setting, adopted this approach and found suggestive effects of exposure to (fictional) news stories in terms of what issues experimental subjects judged to be important.

Agenda setting is normatively important on its own. Sunstein (2001) suggests that the broadcast media serve—or at least ought to serve—the important role of “creating a kind of shared focus of attention for many millions of people” (p. 35). As such, they create shared experiences among a heterogeneous public and expose people to issues and problems they may not have known about or considered previously. Moreover, as Rogers and Dearing (1988) suggest, the media agenda may affect the public’s agenda, which in turn may have consequences for what issues politicians focus on, thus translating into important policy consequences.

Priming

In addition to being important on its own, agenda setting is also important because it relates to priming. Priming refers to the phenomenon that the issues that people judge to be important often become the criteria by which they evaluate politicians (e.g., Krosnick & Kinder, 1990). Thus, any factor affecting public agendas has the potential of affecting politically relevant variables such as presidential approval and the vote. One example of this process concerns George H. W. Bush’s failure to get reelected in 1992, despite having just achieved victory in the first Gulf War. The reason for this, Zaller (1994) argues, is that media attention had shifted from the war to the economy, which was allegedly stalling. As a consequence, the public was primed to evaluate Bush not on the war but on the economy instead, leading to a drop in approval ratings and eventually a failed reelection bid.

Framing

Agenda setting and priming thus concern what issues are on people’s minds. Framing, in contrast, is concerned with the presentation of issues or events and the extent to

which the nature of the presentation affects people’s opinions about the issue. In Druckman’s (2001) words, “A framing effect occurs when in the course of describing an issue or event, a speaker’s emphasis on a subset of potentially relevant considerations causes individuals to focus on these considerations when constructing their opinions” (p. 1042).

In an important piece, Nelson, Clawson, and Oxley (1997) argue that issue frames can affect issue opinions. The authors conducted an experiment surrounding the issue of whether a certain community ought to allow a Ku Klux Klan rally. One group of participants was shown a news clip focusing on freedom of speech considerations (the Ku Klux Klan, however noxious, has a right to express its opinions), while another was shown a clip emphasizing the public order dimension of the rally (i.e., the potential for violence). The results show that those exposed to the free speech frame were much more likely to support allowing the rally than those who watched the public order frame. Thus, public opinion responds to the different ways in which an issue or controversy can be presented.

Factors Limiting Subtle Media Effects

The research summarized previously thus provides evidence that the media can have significant effects. However, much of this work is experimental, and although that enhances its internal validity (i.e., the confidence one can have in terms of causality), it potentially reduces its external validity (the extent to which one can confidently say that the observed effects occur outside of the experimental setting). Fortunately, however, more and more researchers are starting to think deeply about the factors that may limit or encourage media effects. All in all, this represents an effort to lay out the conditions under which media influence is more or less likely, and for whom it is more or less likely.

Zaller’s (1992) work is crucially important here. Most pertinently, he presents the idea that those most likely to be affected by media content are also those least likely to be exposed to it. More specifically, those with low political knowledge—and thus the least ammunition to counter media messages—also tend to be those least interested in politics and thus least likely to watch programs or read materials most likely to contain political information. Those most interested in politics and most likely to receive political media content also tend to be the most partisan and thus—in line with the minimal effects tradition—unlikely to be swayed by media content.

Based on Zaller’s work, one can draw two conclusions. First, the salience of a political issue or event matters greatly. Highly salient political information will reach even those least interested in politics, and since these people are susceptible to media effects, this information will likely affect public opinion, public agendas, and, possibly, approval of politicians. Second, there is a group of people in

the middle of the political-interest spectrum who regularly receive political information but who are more susceptible to media effects than highly interested and knowledgeable people. It is this portion of the public who may routinely be affected by media content.

Druckman (2001) has dedicated a line of research qualifying claims about the prevalence and force of framing. In one study, he finds that the credibility of the media source matters in terms of how effective a frame is. Information presented as coming from the *New York Times*, for example, had much greater effects than the same information presented as coming from the *National Enquirer*. This speaks to the larger point that one ought not speak about media effects generally. Instead, one should think about the effects of certain types of information, presented in certain ways, by a certain source, to a certain audience, and under certain circumstances.

Another important point to consider, especially in terms of framing, is that frames seldom have monopoly status (Nelson & Kinder, 1996). If one party or candidate frames an issue in a certain way, the other party or candidate will likely try to find a different way of framing the issue, which may be equally or more compelling. This then forces people to balance these frames. This is not to say, of course, that any counterframe will eliminate a framing effect and that the numerical balance of frame mentions does not matter. All else equal, a stronger frame will beat a weaker frame, and in case of two equally strong frames, the one uttered more frequently will likely prevail.

Another factor that may temper framing effects is political conversation. This argument is akin to one put forth by Druckman and Nelson (2003). Their point of departure is a perceived shortcoming of most framing studies. Most of these studies, they suggest, expose participants to a stimulus and then have them report their opinions “without any social interaction or access to alternative sources of information” (p. 730). That is, these studies place participants in a social vacuum, prohibiting them from discussing the issue with others. As such, much of the framing literature fails to incorporate the social communication dimension of public opinion formation. The authors attempt to address this weakness in the literature by conducting a series of experiments that test the impact of different kinds of discussion on the direction, magnitude, and persistence of framing effects. The differences in discussion type—that is, the different conditions—pertain to exposure to different perspectives. They find that poststimulus discussions that include only common perspectives have no effect on elite framing, but discussions that do include different perspectives eliminate framing effects.

Miller and Krosnick (2000) investigate precisely what mechanism underlies priming effects. The conventional wisdom used to be that exposure to discussion of a certain issue would simply increase the cognitive accessibility of that issue. In other words, the issue would come to be at the top of people’s minds, thus increasing the likelihood

that they would mention it when asked what issue they deemed most important to the country. The authors propose an alternative mechanism: perceived importance. In an experimental study, they find that perceived importance seems to be more important than accessibility. They argue that this is a normatively encouraging result because simple accessibility implies thoughtless application of a criterion, whereas perceived importance suggests a more thoughtful, deliberative approach to the question of what issue is most important and more relevant to politician approval judgments.

New Media

Over the last few decades, the media landscape has changed dramatically. The most important change is from an old media model of broadcasting to a new media model of narrowcasting. Broadcasting refers to media appealing to the general public and is exemplified by network television, radio, and newspapers. Narrowcasting, made possible by cable, Internet, and satellite radio, is targeted to very specific audiences.

The new media have a number of important characteristics that set them apart from the old media. First, there is great variety in content, both in terms of breadth (the number of topics) and in terms of depth (the amount of information on such topics). Old media had content limitations, imposed by considerations of time and space. The new media have no such limitations. Second, the new media have much greater user control over what information people are exposed to. In the broadcast model, the media decided what information to transmit, and people had very few alternative sources of information to go to. In the new media model, people have much greater capabilities to select their own sources and to dig deeper when they feel it is necessary.

One concern that scholars have expressed about the new media era is that it may amplify already existing disparities in terms of political knowledge and participation (Prior, 2007). One advantage of the broadcast media era was that it was conducive to passive learning. Even people uninterested in politics would likely encounter and absorb political information, because there were few easily available alternative means of entertainment: Once they were watching television, they would watch whatever was on, which sometimes meant they would be exposed to politically relevant information. Thus, while not motivated to learn about politics, people would pick up political information along the way (Zukin & Snyder, 1984).

This has changed radically in the new media era, which is characterized by a diversity of media options, thus allowing people to opt out of the media outlets or programming that offer political information. Baum and Kernell (1999) point to exactly this development, in the context of the rise of cable, to explain the decreasing audiences for presidential

speeches. In the broadcast days, television viewers had nowhere to turn when a presidential speech came on. In the cable era, people have and take the opportunity to change the channel to an outlet that does not broadcast the speech, thus decreasing audience size. As a consequence, according to Baum and Kernell, networks have become more reluctant to grant presidents airtime, since there is a high likelihood that viewers will avoid or move away from the networks when the speech is on.

Although motivation to seek out information was thus less important in the broadcast era, it is crucial in the new media era. Further evidence of this is provided in a study by Tewksbury (2003), who investigates people's online behavior, with specific attention to the extent to which people seek out political information. Although high percentages of people tend to report that they follow political news, Tewksbury argues that these self-report measures are likely inflated because people like to present themselves as good citizens. His approach, then, is to actually track people's web-surfing behavior. In short, he finds that the percentage of people accessing politically relevant content and the percentage of the total page hits represented by politically relevant content are much lower than the self-report measures suggest. Thus, it appears that when given the opportunity to opt out of consuming political information, many people take it.

In another piece, Althaus and Tewksbury (2002) investigate how the agenda-setting power of the media may have changed in the new media era. After all, the old media had a near monopoly on political information, and thus had great potential to highlight certain issues over others. In the new media era, people may follow their own issue interests or avoid politics altogether, thus lessening the agenda-setting power of the media. The authors compared issue priorities of participants who read the paper version of the *New York Times* with those of people who read the online version. The latter, as it turned out, were less affected by the issues that the *New York Times* emphasized. The online format, Althaus and Tewksbury suggest, gives people the opportunity to avoid content that journalists, editors, and politicians want to prioritize. This could be good because it is empowering and may provide incentives for politicians and the media to pay attention to issues that the public is interested in, but it could also be bad because the public may be better off focusing on issues that they are not naturally interested in but that are important nonetheless. And given the results of Tewksbury's earlier study, it is quite likely that rather than attending to other political information, people will attend to nonpolitical content such as entertainment and sports stories.

Even among people who are politically interested, the new media may be problematic. Sunstein (2001), for example, argues that although the Internet and cable offer a great diversity of political information and perspective, politically interested people, who, as also suggested by Zaller (1992), tend to have strong ideological and partisan attachments, will likely engage in selective viewing and

processing of political information. They will gravitate to outlets that are ideologically friendly and, as a consequence, will remain unaware of competing viewpoints, which, in turn, will further strengthen or even radicalize their predispositions. All in all, then, the new media may lead to polarization and will inhibit a proper functioning of the marketplace of ideas.

Although the foregoing is plausible, subsequent research has not entirely confirmed Sunstein's (2001) ominous predictions. Although there are definitely those who behave as Sunstein would predict, most people appear to be aware of opinions on multiple sides of the debate, and most people appear to seek out information from multiple viewpoints (e.g., Iyengar, Hahn, Krosnick, and Walker, 2008). The same study provides evidence that although partisanship may be only weakly related to information searches, people are more likely to seek out information on topics that directly impinge on their lives. When one is a member of a certain issue public (Krosnick, 1990), one is more likely to be informed about that issue.

A final factor that has been routinely found to affect learning behavior, in both the old and new media eras, is general political interest. People with high levels of such interest, who tend to be well educated, are informational omnivores: They seek out and absorb information about a lot of issues and topics. Given this and the easy availability of a wealth of information in the new media era, their informational edge is likely bigger now than it was in the old media age. And since passive learning occurs to a much smaller degree in the new media era, those who are not interested are likely to absorb even less information than before, thus compounding the disparity in information (Prior, 2007).

New Media and Participation

Putnam (2000) attributes a significant part of the blame for an observed decline in civic engagement to the rise of television. Before television, he argues, people were more likely to spend their leisure time in civic organizations or otherwise interacting with others, often leading to connections that encouraged and facilitated civic action down the line. Television changed all this, because it offered easily accessible diversion in the home, thus removing incentives to join organizations and engage in civic projects. This, Putnam suggests, led to an erosion of social connections and all of the publicly oriented benefits thereof.

The new media era poses both increased threats and increased possibilities for political participation, and it is again important to think about this in terms of different consequences for different people. For people who are less politically interested to begin with, the new media era removes opportunities and incentives to venture out of the home, be exposed to political information, and be encouraged to participate in politics, be it by voting, protesting, or attending town meetings. Moreover, lacking such social connections, people are less likely to develop the skills and interests necessary to such political participation.

For people who do have an interest in politics, the Internet provides the opportunity to gain more information, to draw attention to an issue, to recruit other participants, to raise money, and to organize face-to-face meetings with like-minded individuals and opponents. Thus, for these people, the Internet offers yet another outlet and source for their political interest and participation. In other words, the Internet has the potential of creating a political context in which there is a minority of highly informed, highly active, and, potentially, highly partisan participants and a large majority of uninformed, uninterested, and unengaged people (Prior, 2007).

Future Directions

Given the dynamic nature of the media landscape, there is plenty of opportunity and need for new research. One fruitful area concerns the consideration of politicians' strategies in the new media era. What role do candidate websites play, both in terms of getting the message out and in terms of raising funds? How successful are campaigns' attempts to have YouTube clips and other types of clips go viral and to what effect? Along similar lines, what effect do citizen-or interest-group-created election materials have?

It is clear that the new media offer both promise and dangers to widespread public participation in politics. How do we harness this potential, and how do we avoid the dangers? Do social networking sites like Facebook and MySpace necessarily distract people from politics, or can they also inform and mobilize people and spur the type of face-to-face interaction that Putnam (2000) so values?

Finally, legend has it that people who viewed the first presidential debate between Kennedy and Nixon on television believed JFK had won, while those who listened to it on the radio believed Nixon had the edge. Could similar media effects be at work nowadays, benefiting candidates who are particularly new media savvy? Similarly, does the rise of high-definition television have implications for which candidates may be successful?

References and Further Readings

- Althaus, S. L., & Tewksbury, D. (2002). Agenda setting and the "new" news: Patterns of issue importance among readers of the paper and online versions of the *New York Times*. *Communication Research*, 29, 180-207.
- Ansolahehere, S., & Iyengar, S. (1996). *Going negative: How political advertisements shrink and polarize the electorate*. New York: Free Press.
- Bagdikian, B. (1997). *The media monopoly* (5th ed.). Boston: Beacon Press.
- Baum, M., & Kernell, S. (1999). Has cable ended the golden age of presidential television? *American Political Science Review*, 93, 99-114.
- Cohen, B. (1963). *The press and foreign policy*. Princeton, NJ: Princeton University Press.
- Dalton, R. J., Beck, P. A., & Huckfeldt, R. (1998). Partisan cues and the media: Information flows in the 1992 presidential election. *American Political Science Review*, 92, 111-126.
- Dautrich, K., & Deneen, J. N. (1996). Media bias: What journalists and the public say about it. *Public Perspective*, 7, 7-19.
- Druckman, J. N. (2001). On the limits of framing effects: Who can frame? *Journal of Politics*, 63, 1041-1066.
- Druckman, J. N., & Nelson, K. R. (2003). Framing and deliberation: How citizens' conversations limit elite influence. *American Journal of Political Science*, 47, 729-745.
- Erikson, R. S., & Tedin, K. L. (2007). *American public opinion: Its origins, content, and impact*. New York: Pearson Longman.
- Finkel, S. E., & Geer, J. G. (1998). A spot check: Casting doubt on the demobilizing effect of attack advertising. *American Journal of Political Science*, 42, 573-595.
- Franz, M. M., Freedman, P., Goldstein, K., & Ridout, T. N. (2007). *Campaign advertising and American democracy*. Philadelphia: Temple University Press.
- Freedman, P., & Goldstein, K. (1999). Measuring media exposure and the effects of negative campaign ads. *American Journal of Political Science*, 43, 1189-1208.
- Graber, D. A. (2003). The media and democracy: Beyond myths and stereotypes. *Annual Review of Political Science*, 6, 139-160.
- Heinderyckx, F. (1993). Television news programmes in Western Europe: A comparative study. *European Journal of Communication*, 8, 425-450.
- Iyengar, S., Hahn, K. S., Krosnick, J. A., & Walker, J. (2008). Selective exposure to campaign communication: The role of anticipated agreement and issue public membership. *Journal of Politics*, 70, 186-200.
- Iyengar, S., & Kinder, D. R. (1987). *News that matters: Television and American public opinion*. Chicago: University of Chicago Press.
- Iyengar, S., & McGrady, J. A. (2007). *Media politics: A citizen's guide*. New York: W. W. Norton.
- Kernell, S. (1992). *Going public: New strategies for presidential leadership*. Washington, DC: Congressional Quarterly Press.
- Krosnick, J. A. (1990). Government policy and citizen passion: A study of issue publics in contemporary America. *Political Behavior*, 12, 59-92.
- Krosnick, J. A., & Kinder, D. R. (1990). Altering the foundations of support for the president through priming. *American Political Science Review*, 84, 497-512.
- Krüger, U. M. (1996). Tabloidization of information on private television. *Media Perspektiven*, 7, 362-375.
- Lau, R. R., Sigelman, L., Heldman, C., & Babbitt, P. (1999). The effects of negative political advertisements: A meta-analytic assessment. *American Political Science Review*, 93, 851-875.
- Lau, R. R., Sigelman, L., & Rovner, Y. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. *Journal of Politics*, 69, 1176-1209.
- Lazarsfeld, P., Berelson, B., & Gaudet, H. (1948). *The people's choice*. New York: Columbia University Press.
- Lippmann, W. (1922). *Public opinion*. New York: Macmillan.
- Miller, J. M., & Krosnick, J. A. (2000). News media impact on the ingredients of presidential evaluations: Politically knowledgeable citizens are guided by a trusted source. *American Journal of Political Science*, 44, 301-315.

- Nelson, T. E., Clawson, R. A., & Oxley, Z. M. (1997). Media framing of civil liberties conflict and its effect on tolerance. *American Political Science Review*, 91, 567-583.
- Nelson, T. E., & Kinder, D. R. (1996). Issue frames and group centrism in American public opinion. *Journal of Politics*, 58, 1055-1078.
- Patterson, T. E. (1993). *Out of order*. New York: Vintage.
- Patterson, T. E., & Donsbach, W. (1996). News decisions: Journalists as partisan actors. *Political Communication*, 13, 453-468.
- Prior, M. (2007). *Post broadcast democracy: How media choice increases inequality in political involvement and polarizes elections*. Cambridge, UK: Cambridge University Press.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- Rogers, E. M., & Dearing, J. W. (1988). Agenda setting research: Where has it been, where is it going? In J. Anderson (Ed.), *Communication yearbook 11* (pp. 555-594). Newbury Park, CA: Sage.
- Sunstein, C. R. (2001). *Republic.com*. Princeton, NJ: Princeton University Press.
- Tewksbury, D. (2003). What do Americans really want to know? Tracking the behavior of news readers on the Internet. *Journal of Communication*, 53, 694-710.
- Vallone, R., Ross, L., & Lepper, M. (1985). The hostile media phenomenon: Bias perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology*, 49, 577-588.
- Wattenberg, M. P., & Briars, C. L. (1999). Negative campaign advertising: Demobilizer or mobilizer? *American Political Science Review*, 93, 891-899.
- Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge, UK: Cambridge University Press.
- Zaller, J. R. (1994). Elite leadership of mass opinion: New evidence from the Gulf War. In W. L. Bennett & D. L. Paletz (Eds.), *Taken by storm: The media, public opinion, and U.S. foreign policy in the Gulf War* (pp. 186-209). Chicago: University of Chicago Press.
- Zukin, C., & Snyder, R. (1984). Passive learning: When the media environment is the message. *Public Opinion Quarterly*, 48, 629-638.

83

U.S. CONGRESS

TOM MILES

University of North Texas

Study of the U.S. Congress is one of the largest areas in the American politics literature. As the so-called first branch of our three major political institutions, Congress enjoys the power of the purse in that it controls the budget and the appropriation of monies to federal agencies and to the states. Congress is also responsible for passing the laws, or policies, that govern the lives of all citizens of the United States and can even amend the Constitution, which it has done 27 times in the nation's history. As such, the U.S. Congress merits scholarship in its own right.

Yet, as any student of political science knows, Congress does not operate in a vacuum. The Founders sought to balance the federal government by giving shared responsibility to all three branches. In addition to the obvious power of the veto, the president also acts in much more subtle ways to try to gain influence over the policies and budget priorities of Congress. The federal agencies, to which the budget is appropriated, also under the president's purview, then have power over the implementation of legislative policies, which are often ambiguously worded. The courts, with the power of judicial review, frequently deem the acts of Congress unconstitutional, thus overruling the actions of elected officials. Additionally, members of Congress are under pressure from their political parties and their electorates, who push them to enact policies that best accord with their respective wishes. One studying the legislature must necessarily acknowledge these pressures

faced by members of Congress and ask how the pressures affect the behavior of the institution and its 535 individual members.

This chapter looks at the research on these complex and dynamic relationships. Beginning with a brief look at the formative literature, the chapter turns to the research on the effect of public opinion on Congress and vice versa. This is followed by a look at the dominant areas of discourse on the separation-of-powers literature. Finally, the chapter looks at some potential future directions of congressional scholarship and summarizes the discussion.

The Theoretical Background

The empirical study of Congress arguably started with David Mayhew's (1974) *Congress: The Electoral Connection* and Richard Fenno's (1973) *Congressmen in Committees*. Much of what they wrote in this important decade has come to frame the field of congressional research. Focusing on the pressures facing legislators, much of their debate centers on the relative influence of constituents, interest groups, the president, and political parties as determinants of the policy decisions members of Congress make. From there, the field has grown to encompass increasingly complex studies of these dynamic interrelationships. This section looks at these questions as the theoretical base of legislative research.

David Mayhew's (1974) book is one of the most enduring works written about Congress and has become the impetus for much subsequent research. He builds on the influential rational choice theory of Anthony Downs (1957), which argues that humans, being rational, will naturally seek alternatives that further their goals. Members of Congress, being human, must then seek the best policy positions that meet their most immediate and tangible goal: reelection. This explains why congresspersons often seek so-called pork barrel spending that benefits their individual constituencies; it enables them to return to their districts and claim credit for creating jobs and prosperity. All other functions, such as creating policy that serves the broader interest and contributing to the maintenance of the chamber, are secondary to the reelection goal. Serving on important committees and gaining rank within them, for instance, is integral to the process of bolstering their images among the electorate.

This line of analysis is, however, not without its critics. In the 1970s, Richard Fenno (1973, 1978) wrote a series of books highlighting the complex motivations of legislators. Contrary to Mayhew, Fenno argues that members of Congress have other motivations as well, such as gaining power and prominence through high-ranking committee positions and enacting good and lasting public policies that extend in many ways beyond simple reelection goals. In his book, Fenno (1973) follows the activities of several congresspersons and observes that members of Congress in general pass through two phases in their careers: the expansionist and the protectionist. During the first phase, the new member seeks to expand the base of voter support within the district and seldom undertakes risky policy positions. However, once the congressperson has secured the trust of the constituents, there exists a level of confidence that enables him or her to pursue more broadly based policies. The congressperson must still be wary of potential threats, hence protectionism, and must spend a large amount of time in the district engaging in constituent service and cementing ties to the communities, yet he or she still enjoys a greater latitude in engaging the larger needs of the chamber. As one would expect, senators, who face reelection only every 6 years rather than 2 and have larger, more heterogeneous districts, exercise a far less personal touch in their districts.

Both Fenno (1973, 1978) and Miller and Stokes (1963) highlight another important distinction in legislative studies: the difference between a delegate and a trustee relationship between the member and constituents. Under the instructed delegate theory, members are seen as faithful executors of the policy preferences of the voters. A trustee, however, will believe that by electing him or her, the voters entrusted the congressperson with the ability to make decisions both in the interest of the district and in the broader national interest. Much of this distinction must, of course, hinge on how well the congressperson knows the will of the constituency and, as Fenno (1973) points out,

how secure the member feels about future reelection prospects. Although most scholars now acknowledge that both relationships are at work in varying amounts among individual congresspersons, it is still a matter of debate as to which theory is more prominent.

Miller and Stokes's (1963) work also develops a dominant theme in the study of the legislature: the concept of a responsible two-party system. In an influential article, they write about the so-called normative consequences of having a legislature that is more responsive to the demands of its individual constituencies than it is to the demands of national policy in a responsible party system. Certainly, it appears that the Founders did intend for the House of Representatives to be the purveyors of constituent will by making them accountable for their decisions every 2 years at election time. Such legislators will tend to act like 435 individual agents, and it is not clear that party voting will be very important to them. Yet Miller and Stokes argue that Congress does in fact conform to modified party voting with members voting along party lines, except in cases where there is clearly articulated electoral preference—a line that they cross only at their own hazard. They find that party control and electoral accountability are both at work and that the relative influence of each varies by issue area, with legislators being most responsive to constituent demands in the area of civil rights policy.

A similar debate exists as to the role of parties and coalitions in Congress. Reacting to the work of Theodore Lowi (1979), who contends that the rising influence of interest groups who provide valuable campaign funds and the growing power of the executive branch have diminished the role of political parties in Congress, Cox and McCubbins (2007) argue that parties still exert substantial control. If it is the case that the congressional committees are at the beck and call only of executive agencies that dictate policy from above and the organized interests who seek to fulfill their financial goals, as Lowi insists, then no collective action would be achievable in Congress. Similarly, while Cox and McCubbins agree with Mayhew that members of Congress are principally concerned with reelection, they criticize the notion that parties are too internally divided by reelection-seeking individuals to be effective. Instead, they argue that parties in Congress, by using procedural rules and holding members accountable, act as coalitions to organize the membership and make collective action possible. Congresspersons react to their partisan reputation not only in their dealings within the institutional setting but also as a means of organizing issues for their voters who, after all, choose them in partisan elections.

An additional question in the literature is as follows: How responsive are members of Congress to their constituents' will, and how accurately are they able to discern that will? It appears the best way to accomplish this is to spend a great deal of time in their districts, listening to the concerns of their constituents and attending events in the

communities. This grassroots activity has earned members of Congress, particularly in the House, the reputation of being the Tuesday to Thursday Club, trying to achieve as much as possible in the middle of the week in Washington so they can rush back to their districts to do the real work that will enhance their electoral prospects. Perhaps a better nickname, according to Davidson and Oleszek (2000), is the “two Congresses.” There is, they tell us, the textbook Congress of bills and laws, rigidly bound by rules, norms, and procedures. However, there is a second Congress of 535 individuals who realize that their fortunes ride on the perceptions of their voters. This creates a situation in which members must maintain a precarious balance between their two roles.

A very concise and clear explanation of the influence of these forces all competing for influence over the policy decisions of Congress can be found in Jacobson (2001). In this book, he details the crux of the electoral dilemma facing congresspersons—namely that they face the increasing influence of national and party demands but still must maintain an image among the voters in the district as being one of them. Importantly, Jacobson asks the following questions: To whom are the members of Congress most accountable, and whose interest do they most reliably serve? Offering an answer to these questions necessarily entails an understanding of the role of public opinion. The following section addresses this line of inquiry.

Congress as Public Enemy Number One?

A subsequent development in the line of research started by Mayhew (1974) and Fenno (1973) resulted in the study of public perceptions of Congress as a whole. Although the aforementioned literature looks at the motivations of Congress’s members, several scholars have chosen to look through the opposite lens and ask how the actions of Congress translate into levels of approval or disapproval among the broader public. In particular, why does Congress have such a low level of support, typically ranging between 10% and 20%, while constituencies tend to overwhelmingly approve of their individual congressperson on the order of 70% or better? As Mayhew and Fenno indicate, members also enjoy a significant incumbency advantage and, with the exception of certain realigning elections, are returned to office close to 90% of the time. This occurs despite a host of scandals and overall levels of dissatisfaction that have plagued Congress in recent years. This section next addresses this growing area of research.

There are a number of reasons for the public dissatisfaction that has branded Congress as public enemy number one in American politics. Research indicates that Congress is typically perceived as being too characterized by bickering and political infighting to get the nation’s business done. This could help explain why Congress has the lowest level of public approval when compared with the president

and the Supreme Court. Congress is also thought to be the most out of step ideologically with the American public. These contentions are backed up by polls conducted by Hibbing and Theiss-Morse (1995) in their book. Surprisingly, they find that most Americans clearly perceive that Congress is the most powerful of the three branches, with an institutional authority that outstrips even that of the more visible office of the president. In this atmosphere, it is reasonable to assume that when Americans are dissatisfied with the pace of governmental response to a pressing national problem, it will be Congress that absorbs much of the blame. Certainly, presidents are able to capitalize on this perception and frequently blame Congress for the bad fortunes of the nation.

Further, Congress is the branch most believed to be a part of the Washington system. Many Americans see politics as increasingly responsive to the wishes of special interests, greed, and corruption. Certainly, since Congress lacks the term limit of the president, many members enjoy long tenures in their respective houses. Consequently, they are likely to be seen as professional politicians firmly ensconced in the nasty world of Washington dealings. A series of high-profile corruption charges including taking money, personal vacations, and bribes from lobbying groups led to several attempts to reform congressional ethics in the 1990s. Despite these reforms, many Americans are likely to view members of Congress as being more driven by personal gain and the perks of office than by a genuine desire to act in the public interest.

Tellingly, Hibbing and Theiss-Morse (1995) find that most Americans see Congress as a distinct part of the two political systems that inhabit American politics. Although the people interviewed in their study tended to view the legitimacy of the presidency as an institution separately from the current president, Congress does not enjoy this level of public discernment. Rather, the legitimacy of Congress is intimately tied to the actions of its current members, a case in which the actions of a few members can poison the well of public satisfaction with the institution as a whole. This tendency is exacerbated, they tell us, by the fact that few Americans are politically engaged, which leads to lower levels of sophistication. Studies have shown (Born, 1990) that persons with lower levels of political knowledge and sophistication have a more difficult time disentangling the actions of individual members and the performance of Congress as a whole.

All of this takes place amid an atmosphere of apathy and general dissatisfaction with American politics. And since Americans generally pay little attention to politics, more of them are likely to hold ambiguous political attitudes. As Hibbing and Theiss-Morse (1995) point out, the American public in general is very critical of pork barrel politics as a vehicle for wasteful spending and excessive government programs. Yet when their own congressperson brings jobs and money to the district, he or she is perceived very favorably in the district. What is more, although much of the

American electorate feels disenchanted by what it sees as an increasingly polarized political system that is either too far to the left or right, it often acts in ways that perpetuate extremes and hyperbole in the political rhetoric. The result is a public that is generally displeased with the policy options being presented by either party (see Fiorina, 1992) and thus disengaged from the political process.

Despite the perennial calls to throw them all out at the next election, this seldom happens. Despite scandals and perceived inefficiency and waste, members of Congress continue to enjoy a significant incumbency advantage. So scholars naturally wish to know what conditions are necessary to turn enough of the electorate against an incumbent and toward a newcomer. Kinder and Kiewiet (1979) placed this question prominently into the scholarly debate in their inquiry into the role of economic conditions in congressional elections. Kinder and Kiewiet find, perhaps surprisingly, that the voters they studied responded less to individual economic grievances and focused more on the prospects for overall national economic growth. The study, however, has perhaps raised more questions than it answered and has spawned significant debate. These are among the questions: How do we account for those voters who participate irregularly—what brings them to the polls—and what is the linkage that ties individual evaluations to broader political perceptions? And, more important, what is the tipping point at which an election becomes critical and results in a large electoral realignment?

Certainly, large shifts in public opinion can affect significant electoral realignments such as were seen in the 2006 and 2008 elections. The combined effect of two wars, in Iraq and Afghanistan, and a global recession cost Republicans and resulted in large Democratic majorities in both the House and Senate. But absent extreme conditions, it remains unclear if voters in congressional elections vote with their pocketbooks—that is, strictly on retrospective perceptions of how Congress is affecting their personal financial well-being—or vote sociotropically (as Kinder and Kiewiet, 1979, claim), with an eye to overall economic conditions and general welfare. Or do large groups of the electorate focus primarily on social issues they care about, like the environment or abortion, and evaluate their congressperson mostly on these evaluative bases? And further, are these evaluations of Congress substantively different from those of other elected officials, most notably the president? These are large questions for future research and dovetail with the even wider literature on voting behavior.

Although all of these phenomena unquestionably characterize the uneasy relationship between Congress and public opinion, how do they affect the performance of Congress in terms of its role in forming public policy? This chapter has hinted at consequences of shared power among the institutions that make up the national system, but in fully evaluating Congress, it is necessary to flesh out these interrelations. The next section looks at some of these areas of shared responsibility.

Policy Implications

The aforementioned literature bespeaks the fact that Congress is a very complex amalgam of individuals with often disparate goals. Members of Congress must balance the needs of their constituents and the prospect of facing an impending reelection campaign against the needs and goals of the government at large. Additionally, Congress must share the responsibility for shaping public policy with the president, who often calls on Congress to enact certain types of legislation that accord with his policy goals, and with the courts and bureaucracies that are tasked with implementing those policies. In terms of policy implications, some of the more important questions are as follows: How much does divided government matter in the shaping of policy? How large is the president's role? To what extent do courts and executive agencies determine the eventual output of the policy-making process in implementing law? The following section highlights the literature on each of these important issues.

The Question of Divided Government

The previous section discussed governmental gridlock, and it is still a matter of debate in congressional literature as to the degree to which the president and Congress being controlled by opposing parties stymies legislation. Edwards, Barrett, and Peake (1997) nicely summarize much of this research. They take the arguments offered by Mayhew (1991) and Kernell (1991) as the two predominant sides of the debate. Mayhew argues that gridlock is not a serious problem since the same amount of legislation is passed when Congress and the executive are of different parties as when they are unified under one party. Kernell, on the contrary, finds that divided government leads to inevitable conflict, which slows down the passage of important legislation. Clearly, this is a question that is far from settled.

What Edwards et al. (1997) find, however, is interesting and supportive of the position argued by Kernell (1991). They categorize bills according to importance and find that passage of important legislation is more likely to be slowed down substantially under conditions of divided government. Further, presidents are more likely to oppose more bills than when their party controls Congress. Mayhew's (1991) finding that the volume of bills passed is the same in either case indicates that the workload for Congress is the same but that less significant legislation is being passed. Although these findings are important in their own right, Edwards et al. insist there is still more work to be done. For example, little research has been done on the dilution of bills as a result of interbranch bargaining during times of divided government. According to this hypothesis, interaction between Congress and the president may result in watered-down versions of key legislation that will be less

likely to provoke a presidential veto. The extent to which this occurs is still very much a matter for debate.

Agenda Setting

Closely related to the effect of divided government is the process of presidential bargaining with Congress, which is known as agenda setting. This is one of the more widely explored areas of congressional politics and encompasses a broad literature. Yet despite all that has been written, questions remain as to how much sway the executive branch holds over the legislative. All indications are that it is a very complicated relationship and varies between presidential administrations, dependent on the president's bargaining skills, ideological congruence between the Congress and the president, and how skillfully the president can enlist public sentiment through the media and public speeches.

Although, as noted, Congress controls the purse strings and must approve all spending, it is the president's responsibility to recommend an annual budget. This gives the president the ability to try to bargain with Congress over what he or she believes to be the important spending priorities in the upcoming annual fiscal budget. Similarly, each year in January, the president is responsible for delivering the State of the Union address on the floor of Congress. These are notable instances in which the president has a large and very tangible opportunity to influence the domestic agenda or at least put himself into a bargaining position with Congress. The process does not stop there; the president will also use major speeches and public appearances throughout the year to keep pressure on Congress by continually calling attention to these issues. Conversely, for potential bills that the president does not favor, he or she will use speeches as an attempt to drum up negative public sentiment and will, at times, even threaten a veto.

How successful these tactics are varies between presidents and issues. When the president stakes out an issue position and attempts to enlist public support to influence Congress, the consequences of failure can be high and must be weighed against the potential for success. If a president takes a position on a popular legislative initiative and fails, he risks being seen as weak and ineffectual. However, if the president does not take a position on an important issue, he may perhaps also be seen as weak—a so-called lame duck who abrogates too much power to Capitol Hill. Important research on this executive-legislative give-and-take is to be found in the works of Light (1999), Bond and Fleisher (1990), and Eshbaugh-Soha (2005). Their analyses of this complicated process reveal that much of the president's success comes from the way the yearly agenda is packaged and how able the president is to choose issues that have high probabilities of success in Congress. Also, as mentioned previously, the president is more likely to win when his party controls Congress and

he has party factions favorable to his policy positions in ranking committee positions. What is more, as Eshbaugh-Soha points out, the type of policy and how aggressive it is often dictate how much pressure the president can afford to exert. By pushing for too much too soon, a president risks failure even under conditions of unified government.

Congress and the Courts

Even if Congress is successful in mustering majorities and passing legislation that the president signs into law, all of that work can be undone by a lawsuit that works its way into the federal court system. District and circuit courts of appeal, most notably the U.S. Supreme Court, can and do use the power of judicial review afforded to them by the famous *Marbury v. Madison* (1803) case to declare acts of Congress unconstitutional. When this happens, Congress must then choose either to abandon the issue or to pass an attenuated piece of legislation that the members believe will avoid invoking the courts again. And the relationship between the courts and Congress extends beyond the potential for judicial review. The Senate must confirm presidential nominations to vacant federal benches, and Congress holds the power to impeach and remove judges. This sets up a complicated interplay between the legislature and judiciary, who must each weight their goals against the potential actions of the other branch.

In their seminal work on the Supreme Court, Segal and Spaeth (2002) insist that Supreme Court justices are not uninvolved in politics but are, rather, firmly entrenched in the policy-making process. They argue that Supreme Court justices take advantage of the fact that legislation is often ambiguous or vaguely worded in order to make decisions that further their own political goals. This fact puts them at odds with Congress since both branches are pursuing often different ideological and policy agendas. If Segal and Spaeth are correct and justices are concerned more with taking political sides than with merely dispassionately dispensing justice, then one potentially must view this system of checks and balances between the two branches as a source of increasing institutional conflict. Important work is current being undertaken to explore the depth of this conflictual relationship.

Congress and the Bureaucracy

Since it is the duty of Congress both to appropriate money to the various bureaucratic administrative agencies and to oversee their operation, it is not surprising that members of Congress are often perceived by the public as shirking the duty of oversight. When the media publicize stories of bureaucratic waste and inefficiency, it is often Congress that absorbs much of the blame. And this fact may go partway to explaining why Congress is seen as public enemy number one in American politics. Presidential

campaigns often fan the flames of public dissatisfaction by heaping blame interchangeably on what they characterize as a bloated, inefficient federal government and an equally untrustworthy Congress.

McCubbins and Schwartz (1984) insist that this process is much more subtle and nuanced than it appears at first blush. Recalling Mayhew's theoretical argument that members of Congress will seek to claim credit for programs that benefit their constituencies and thus their reelection chances, they take great care in how they allocate their resources. Therefore, there is little motivation for members to undertake the more resource-costly police patrol method of overseeing bureaucracy on a daily basis. They instead prefer the so-called fire alarm method of oversight, which calls on them to act when there is reason for doing so and in instances when they can claim credit for having acted in their constituents' best interests. The fact that Congress controls appropriations for executive agencies, on the other hand, gives them a much greater degree of control than is apparent to less-attentive observers. For this reason, McCubbins and Schwartz prefer to view executive agencies as agents of the committees and subcommittees who are responsible for the proportion of the budget that is allocated to the bureaucracies they oversee. A more detailed account of the budgetary and oversight functions can be found in Stillman's (1996) *The American Bureaucracy*. These valuable insights bolster the idea that oversight, like much else connected with the Congress, is more complicated than a cursory examination will reveal.

Congress and Interest Groups

An additional area of concern is the influence of organized interest groups on legislative output. Ever since David Truman's (1971) influential book on interest groups, there has been a growing literature on the sway that social and political lobbies have over the members of Congress. It is well-known that interest groups engage in direct lobbying efforts, providing information and resources to members of Congress, and indirect or social lobbying, which can include hosting fund-raising events. Although a series of campaign finance and lobbying reform acts since the 1990s have sought to curtail the use of lavish gifts and vacations as avenues for lobbyists to gain undue influence over Congress, questions remain over the extent to which interest groups shape the public agenda. Organized interests are even known to author, in whole or part, the language of bills that make it to the floors of the House and Senate.

Some scholars prefer to see the interactions between special interest groups, congressional committees, and the bureaucracies they fund as so-called iron triangles (Knott & Miller, 1987; Truman, 1971). The triangle imagery is meant to convey the idea of a reciprocal system in which benefits flow among the three apexes. Lobbies provide

electoral support to members of Congress, who provide funding and political support to bureaucratic agencies, who in turn provide special favors to interest groups in a self-perpetuating cycle. Other scholars characterize congressional committees and subcommittees and the areas of policy specialization on which interest groups focus their efforts as an unavoidable consequence of administering a huge national system. Members of Congress do not have time to focus on all areas of public policy, so they must become specialists in more specific areas—hence, the congressional committee. The same limited resources of time and knowledge make it necessary for them to rely on information sources outside their limited staffs. Interest groups readily provide this information, often in distilled form. And since interest groups represent the aggregate of citizens to whom these policies are important, this information provides a shortcut to discerning the public will.

Regardless of the value of the information provided, many will continue to see the role of interest groups as an indication of a Washington system run amok. Many are led to question the accuracy of information provided with the goal of currying influence and political or monetary gain. There is similar debate among scholars as to the value of the information provided by interest groups to the academic world. For instance, the liberal Americans for Democratic Action and the conservative American Conservative Union provide scores on individual legislators (0–100) on how well or poorly they voted on liberal or conservative issues, respectively. Some scholars have used these scores as measures of ideological leaning among members of Congress. Other scholars insist that these scores will be slanted by the interest groups' own ideological leanings. Thus, there is little consensus as to the accuracy of these scores. Similarly, the question of how well interest groups represent the public will as opposed to the moneyed minority of citizens that fund them remains unanswered.

Future Directions

The previous section already touched on some open questions in congressional studies. Clearly, many issues remain for legislative scholars. Perusing the current literature in the field reminds one that this is a very complex and dynamic field. Just as other areas of political science have embraced interdisciplinary perspectives, so too has legislative politics been enriched by findings in other fields. In particular, behavioral explanations for the decisions of members of Congress have been offered by scholars in psychology that are enlarging the scope of the debate on party and policy. One of the important questions being asked is whether Congress is also swaying to the purported trend in American politics toward a more partisan orientation.

One of the implications of Mayhew's argument that members of Congress are single-minded seekers of reelection is

that policy itself is not as important as taking a position on an issue that will please as much of one's electorate as possible. Yet in a seminal article in social psychology, Geoffrey Cohen (2003) demonstrates that group pressure from a participant's self-identified partisan cohort (Democrat or Republican) contributed significantly to support for a welfare-policy initiative. When a policy proposal for a work-training program, for instance, was proposed by a Republican, the study participants identified as Republican overwhelmingly supported it and tended to regard it as a conservative proposal by definition. The same was true for Democrats. This occurred despite the fact that there was no difference in the two proposed bills except for the sponsor. Interestingly, the study's subjects denied having been influenced by their political party. Cohen refers to this as the effect of social identity—people are unconsciously constrained in their choices by the types of groups they identify with. It remains to be seen if this same type of group pressure applies to members of the House or Senate, but given that Congress votes overwhelmingly along party lines, it may indeed prove to be the case.

Much research continues to look at the influence of public opinion on congressional behavior. New research is aimed at discerning the interrelation of levels of public support among institutions—Congress, the Supreme Court, and the presidency—over time (see Bailey, 2007). Bailey insists that levels of support among the branches are actually more similar than previously estimated, with the disparity being accounted for by inadequate means of measurement. A separate but similar line of inquiry seeks to identify the effect of information on levels of approval for Congress. In a manner similar to Born (1990), Mondak, Carmines, Huckfeldt, Mitchell, and Schraufnagel (2007) find that higher levels of political knowledge translate into more specific evaluations of Congress whereas lower sophistication tends to lead to more heuristic judgments. These heuristic judgments will tend to be more susceptible to the influence of mass media and political actors that readily paint Congress in a negative light. Much research indicates that the availability of information and its temporal proximity have a great effect on the importance afforded to it unconsciously as the brain processes information.

Political science will continue to benefit from these enlarging perspectives. Research into the cognitive processes that inform public opinion has barely begun and may prove to add substantially to our knowledge of how the information environment interfaces with perception to shape political judgments. Similarly, scholars continue to probe the effect of evaluations of other institutions on government in general and how these translate into support for Congress. As previously mentioned, Congress does not operate in a vacuum, and scholars continue to investigate the cumulative effects of divided government and shared power among the branches. Although this chapter has not specifically mentioned the media's role in shaping Congress's policy priorities, a large body of literature aims

at identifying these important linkages. See Chapter 82, titled "Media and Politics," for that discussion.

Conclusion

What should a 21st-century student of legislative politics know about this complex field of study? Clearly, there are a great many avenues for research, and this chapter has only just brushed on them. Facing an ever-enlarging body of literature and an increasing reliance on the insights being offered by other disciplines, such as psychology, many scholars choose to specialize in only a very narrow area of analysis—agenda setting or the effect of public opinion, for example. Additionally, as the issues faced by Congress become larger and more complex, as a growing population with more diverse interests dictates, so too will the job of studying Congress become more complicated. With this complexity, however, comes the prospect of a seemingly endless range of research possibilities. Study of the legislature is one of the richest areas of American political science.

As such, this chapter is able to offer only a brief look at the many faces of congressional research, which, to summarize, started with the seminal theoretical works and traced their respective studies of the electoral and often competing national pressures that bear down on members of Congress. Next, this chapter looked at the literature on the roles of parties and coalitions in shaping the actions of Congress. From there, the chapter turned to the significant public opinion challenges faced by all legislators as members of an institution that is generally perceived to be corrupt and to be failing in its job of acting in the public interest. The policy implications gleaned from this theoretical base naturally lead to the question of Congress's efficacy in enacting policy. This naturally entails looking at the legislature in its dealings with the other branches of government with which it shares power. The chapter examined this complex question by briefly looking at the literature on presidential agenda setting, the consequences of divided government, and interactions with the federal bureaucracy and the judicial system. Finally, the chapter looked at some potential areas of future research.

These, then, are some of the major themes in legislative study. This brief look, however, does not exhaust the field. What this chapter has sought to do is not only to overview the complexity of congressional research but also to highlight the many competing interests that vie for the attention and favor of Congress. As the first branch of government, the legislature has the huge responsibility not only of controlling the federal budget, but also of making the laws that govern all the states and citizens of the nation. Yet in the Madisonian tradition of the founding and Constitution of the United States, Congress must also share this power to ensure that its sway over national policy is checked and restrained. For all of these reasons, Congress cannot be seen as a simple institution, nor can it

be seen as a mere collection of 535 members with individual constituencies and agendas. It is a lawmaking body characterized by complexity, competition, and compromise, and it is this very mix of interactions that makes Congress so interesting and challenging to study.

References and Further Readings

- Arnold, R. D. (1990). *The logic of congressional action*. New Haven, CT: Yale University Press.
- Bailey, M. A. (2007). Comparable preference estimates across time and institutions for the court, Congress, and the presidency. *American Journal of Political Science*, 51, 433–448.
- Bond, J. R., & Fleisher, R. (1990). *The president in the legislative arena*. Chicago: University of Chicago Press.
- Born, R. (1990). The shared fortunes of Congress and congressmen. *Journal of Politics*, 52, 1223–1241.
- Cigler, A. J., & Loomis, B. A. (Eds.). (1998). *Interest group politics*. Washington, DC: CQ Press.
- Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85, 808–822.
- Cox, G. W., & McCubbins, M. D. (2007). *Legislative leviathan: Party government in the House* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Davidson, R. H., & Oleszek, W. J. (2000). *Congress and its members*. Washington, DC: CQ Press.
- Dodd, L. C. (1977). Congress and the quest for power. In L. C. Dodd & B. I. Oppenheimer (Eds.), *Congress reconsidered* (pp. 361–372). New York: Praeger.
- Downs, A. (1957). *An economic theory of democracy*. Boston: Addison Wesley.
- Edwards, G. C., III. (1980). *Presidential influence in Congress*. San Francisco: W. H. Freeman.
- Edwards, G. C., III, Barrett, A., & Peake, J. (1997). The legislative impact of divided government. *American Journal of Political Science*, 41, 545–563.
- Eshbaugh Soha, M. (2005). The politics of presidential agendas. *Political Research Quarterly*, 58, 257–268.
- Fenno, R. F. (1973). *Congressmen in committees*. Boston: Little, Brown.
- Fenno, R. F. (1978). *Home style: House members in their districts*. Boston: Little, Brown.
- Fiorina, M. P. (1974). *Representatives, roll calls, and constituencies*. Lexington, MA: Heath.
- Fiorina, M. P. (1992). *Divided government*. New York: MacMillan.
- Hibbing, J. R., & Theiss Morse, E. (1995). *Congress as public enemy: Public attitudes toward American political institutions*. Cambridge, UK: Cambridge University Press.
- Jacobson, G. C. (2001). *The politics of congressional elections*. New York: Longman.
- Kernell, S. (1991). *Going public: New strategies of presidential leadership*. Washington, DC: CQ Press.
- Kinder, D. R., & Kiewiet, D. R. (1979). Economic discontent and political behavior: The role of personal grievances and collective economic judgments in congressional voting. *American Journal of Political Science*, 23, 495–527.
- Kingdon, J. W. (1981). *Congressmen's voting decisions*. New York: Harper & Row.
- Knott, J. H., & Miller, G. J. (1987). *Reforming bureaucracy*. New York: Prentice Hall.
- Light, P. (1999). *The president's agenda: Domestic policy choice from Kennedy to Clinton*. Baltimore: Johns Hopkins University Press.
- Lowi, T. J. (1979). *The end of liberalism: The second republic of the United States*. New York: W. W. Norton.
- Mann, T. E., & Ornstein, N. J. (2006). *The broken branch: How Congress is failing America and how to get it back on track*. Oxford, UK: Oxford University Press.
- Marbury v. Madison, 5 U.S. 137 (1803).
- Matthews, D. R. (1960). *U.S. senators and their world*. Chapel Hill: University of North Carolina Press.
- Mayhew, D. R. (1974). *Congress: The electoral connection*. New Haven, CT: Yale University Press.
- Mayhew, D. R. (1991). *Divided we govern: Party control, law making, and investigations, 1946–1990*. New Haven, CT: Yale University Press.
- McCubbins, M. D., & Schwartz, T. (1984). Congressional oversight overlooked: Police patrols versus fire alarms. *American Journal of Political Science*, 28, 165–179.
- Miller, W. E., & Stokes, D. E. (1963). Constituency influence in Congress. *American Political Science Review*, 57, 45–56.
- Mondak, J. J., Carmines, E. G., Huckfeldt, R., Mitchell, D. G., & Schraufnagel, S. (2007). Does familiarity breed contempt? The impact of information bias on mass attitudes toward Congress. *American Journal of Political Science*, 51, 34–48.
- Polsby, N. W. (1968). The institutionalization of the U.S. House of Representatives. *American Political Science Review*, 68, 144–168.
- Segal, J. A., & Spaeth, H. J. (2002). *The Supreme Court and the attitudinal model revisited*. New York: Cambridge University Press.
- Stillman, R., II. (1996). *The American bureaucracy: The core of American government*. Chicago: Nelson Hall.
- Sundquist, J. L. (1981). *The decline and resurgence of Congress*. Washington, DC: Brookings Institution.
- Thompson, D. F. (1995). *Ethics in Congress: From individual to institutional corruption*. Washington, DC: Brookings Institution.
- Truman, D. B. (1971). *The governmental process, political interests, and public opinion*. New York: Knopf.

84

THE PRESIDENCY

REBECCA E. DEEN

University of Texas at Arlington

The phrase *the president of the United States* often conjures grand and expansive titles: leader of the free world or most powerful man in the United States. These labels belie the truth that the president's power is nuanced and derived from the Constitution, from congressional statute, and from aggrandizement by individual presidents. This chapter explores the presidency and the executive branch, first examining the various theoretical perspectives on the sources of presidential and executive branch power. In this section, different academic disciplines that help shed light on this institution are also considered. The chapter next turns to what political scientists have learned about the president with regard to governing and policy making. Finally, the chapter concludes with thoughts about the direction of the study of the presidency.

Theory

In many ways, the powers of the presidency are contradictions born out of the constitutional debates among the framers. Informed by their experiences of tyranny under King George, many of the delegates to the Constitutional Convention wanted no executive at all, or at least one with very limited powers. It is striking that despite the clear deficiencies of the Articles of Confederation (which had no executive branch), discussions of their failings did not include the absence of a chief executive. One of the pressing

debates of the convention was whether to have a singular or plural executive and whether the powers and responsibilities invested in the executive branch should be divided across more than one office. This option would have weakened the presidency, thus allaying the fears of those who saw danger in a powerful chief executive. Indeed, some have argued that the term *president* was chosen because it was an innocuous title, likened to a presiding officer who would exercise little independent power (McDonald, 1994).

A hallmark of our republic is the separation, dispersion, and sharing of power among the three branches of government. An examination of the U.S. Constitution illustrates this. Although the president has the power to command the military, to enter into treaties, and to make executive branch appointments, those powers are constrained by Congress's ability to declare war and the Senate's ratification, advice, and consent roles regarding treaties and appointments, respectively. The framers spent half of the relatively short article on the presidency (1,023 words as compared with 2,267 words in Article I on the legislature) discussing how the president would be selected for the office and when a president could be removed from it.

This description of the office probably sounds foreign to a casual observer of the contemporary U.S. presidency. The office and the public's expectations of it have grown substantially since the nation's early days. This growth can be attributed to the aggrandizement by individual presidents,

by ceding of power by congressional statute, and by the post–New Deal expansion of the role of government.

Though the modern presidency is marked from the administration of Franklin Delano Roosevelt, presidents before him (see Table 84.1) articulated their views on presidential power. Some, like William Howard Taft (1909–1913), viewed the president’s power to be strictly derived from Article II of the Constitution. In his view, the executive’s power was limited only to what was prescribed by the Constitution (Taft, 1916). At the other extreme was President Richard Nixon (1968–1974), who famously remarked in an interview with David Frost on May 20, 1977, “When the president does it, that means it’s not illegal” (Winther, 2008).

Other presidents, such as Theodore Roosevelt (1901–1909), have taken a moderate view, arguing that presidents have a duty to take any actions that are in the country’s

interests, so long as they are not explicitly prohibited by the Constitution (Roosevelt, 1913).

Franklin D. Roosevelt’s administration, faced with the challenges of the Great Depression and World War II, presided over an enormous growth spurt of the federal government that substantially increased the executive branch scope in the lives of ordinary Americans. Government helped to create a permanent web of social and economic programs (unemployment benefits, social security, and aid to the poor) that the executive branch would administer. Combined with congressional action in 1921 to create the Bureau of the Budget, which allowed the president to set the parameters for budget debates with Congress, this era saw a substantial growth of the executive branch. Roosevelt’s administration has become the benchmark by which future presidents’ productivity has been judged.

In this era of the modern president, political scientists generally have pointed to the extraconstitutional powers of

<i>U.S. Presidents in Chronological Order</i>	
1. George Washington	1789 1797
2. John Adams	1797 1801
3. Thomas Jefferson	1801 1809
4. James Madison	1809 1817
5. James Monroe	1817 1825
6. John Quincy Adams	1825 1829
7. Andrew Jackson	1829 1837
8. Martin Van Buren	1837 1841
9. William Henry Harrison	1841
10. John Tyler	1841 1845
11. James Knox Polk	1845 1849
12. Zachary Taylor	1849 1850
13. Millard Fillmore	1850 1853
14. Franklin Pierce	1853 1857
15. James Buchanan	1857 1861
16. Abraham Lincoln	1861 1865
17. Andrew Johnson	1865 1869
18. Ulysses S. Grant	1869 1877
19. Rutherford B. Hayes	1877 1881
20. James Garfield	1881
21. Chester A. Arthur	1881 1885
22. Grover Cleveland	1885 1889

23. Benjamin Harrison	1889 1893
24. Grover Cleveland	1893 1897
25. William McKinley	1897 1901
26. Theodore Roosevelt	1901 1909
27. William Howard Taft	1909 1913
28. Woodrow Wilson	1913 1921
29. Warren Harding	1921 1923
30. Calvin Coolidge	1923 1929
31. Herbert Hoover	1929 1933
32. Franklin D. Roosevelt	1933 1945
33. Harry S. Truman	1945 1953
34. Dwight D. Eisenhower	1953 1961
35. John F. Kennedy	1961 1963
36. Lyndon Johnson	1963 1969
37. Richard Nixon	1969 1974
38. Gerald Ford	1974 1977
39. James Carter	1977 1981
40. Ronald Reagan	1981 1989
41. George H. W. Bush	1989 1993
42. William J. Clinton	1993 2001
43. George W. Bush	2001 2009
44. Barack H. Obama	2009

Table 84.1 U.S. Presidents

SOURCE: White House (n.d.).

the president as his most significant. Arguably the most influential book on the presidency, Richard Neustadt's (1960) *Presidential Power* argues that the power of the president is the power to persuade. Neustadt's hypothesis is that presidents can greatly increase the likelihood of policy success by attending to their professional reputations with other political elites in Washington, D.C., and to their public prestige.

Prior to Neustadt's (1960) work, political scientists largely focused on the limited constitutional powers of the president in trying to understand the office. It is not hyperbole to say that Neustadt fundamentally changed the way scholars think about the office. His work illuminated the path to a number of rich veins of research, for example of congressional–presidential bargaining.

Another perspective that focuses on the personal presidency draws on political psychology. There are a number of approaches. Psychobiography examines biographical information about the president to develop a psychological profile to predict his behavior. One of the most well-known (and criticized) of these approaches is *The Presidential Character: Predicting Performance in the White House* by James David Barber (1972). Barber argued that characteristics of the president's core personality, such as his style of doing the tasks of his job, his worldview about the great issues of his day, and his character or approach to life, combine to influence his being the president. Most of the attention and criticism of his argument has been focused, though, on his typology of presidential character. Barber finds two primary dimensions: the level of energy a president brings to life (active or passive) and the extent to which a president enjoys the job (positive or negative). These two dimensions produce four dominant types: active positive, active negative, passive positive, and passive negative. Barber's typologies have been criticized as being overly simplistic and difficult to measure consistently.

Contrast these presidency-focused heuristics with those that examine the institution or historical epoch of the presidency. Neoinstitutionalism focuses on the resources for and constraints on all presidents. This school of thought takes the perspective that all presidents face these challenges and focuses on these factors as explanations for presidential behavior.

One such work emphasizes the moment in which presidents find themselves. Stephen Skowronek (1997) argues that presidents are constrained by the relationships their policies have to the dominant partisan regimes. He suggests that presidential possibilities and achievements must be seen through the lens of history, what he calls political time. Political time is a function of partisan eras. A dominant partisan regime can be vulnerable to change, or it can be resilient. Similarly, a president might stand in opposition to the regime or be affiliated with it. Combine these two dimensions, and one understands the possibilities for leadership that open for a president, regardless of his or her own personal qualities. For example, a president who stands in opposition to a vulnerable regime (Thomas Jefferson, Abraham Lincoln, or Franklin Roosevelt) or one who is

affiliated with a dominant one (Teddy Roosevelt or Lyndon Johnson) has *prima facie* greater opportunities than one who is opposed to a dominate regime (Richard Nixon) or is affiliated with a vulnerable one (Jimmy Carter).

Applications and Empirical Evidence

The race for the presidency, through the partisan nomination season and the general election, is a marathon. The victorious president-elect would be forgiven for the understandable urge to take a long vacation. However, as presidents-elect and their staffs well know, the weeks immediately following the election are a period where crucial decisions are made. These staffing and organizational choices, coupled with the president-elect's ability to marshal the perception of a mandate, can set the stage for a successful first few months in office.

The Transition

All presidents wish to claim a mandate, though the concept is an illusive one. Electoral landslides (Johnson over Goldwater in 1964, Nixon over McGovern in 1972, or Reagan over Carter in 1980) are seen as clear signals from the electorate. These decisive electoral outcomes, however, are relatively rare. Often, a mandate is claimed when the electoral margin of victory is combined with a statement on the previous administration. Take, for example, Reagan's defeat of Carter, a sitting president. Reagan not only won 51% of the popular vote to Carter's 41% (a third-party candidate accounted for the remaining percentage), but also articulated clearly a vision for the country that was a definite departure from the status quo.

There is relatively little an incoming president can do to affect his mandate. The president does have, however, control over early staffing choices. These choices, if made well, will set him or her up for greater likelihood of success in the critical first 100 days of his or her administration. A president not only must make critical and highly public choices about cabinet appointments but also must make choices about personal staff.

Staffing the Cabinet

The president nominates his or her cabinet (mostly composed of those individuals heading the executive branch agencies such as the Departments of State, Justice, Defense, Commerce, and Interior), though each individual must be confirmed by the Senate. When the Senate is controlled by the party in opposition to the president (this is called *divided party government*), the calculus for choosing cabinet nominees is different than when the president's party controls both the White House and one or both chambers of Congress (called *unified party government*).

The vetting process, investigating to discover information about the nominee's personal or professional life, is

always extensive and sometimes begins even before the conclusion of the election. However, the context of the vetting process is markedly different under divided party government when the outcome of the confirmation process is largely in the hands of the president's partisan opposition. Often, the extent to which opposition-party senators feel free to criticize vociferously a presidential nominee is a function of the president's perceived electoral mandate. A nominee is most likely to face strong opposition when negative information comes to light about his or her personal life. For example, President Clinton nominated Zoe Baird to be attorney general. Ms. Baird withdrew her nomination, however, after information surfaced that she had neglected to pay social security taxes for a former household employee. Similarly, President Barack Obama nominated former Congressman Tom Daschle for Secretary of Health and Human Services but withdrew his nomination when an unpaid tax bill was discovered. Failed nominations usually follow this pattern, with the nomination withdrawn prior to a vote being taken in the Senate. The reasons for this are sensible; a nominee rejected by the Senate is politically damaging for the fledgling administration—as was the case of President George H. W. Bush's choice for secretary of defense, John Tower, or President Dwight Eisenhower's choice for secretary of commerce, Lewis L. Strauss.

Staffing the White House

The choices presidents make for their staff, though less public, are just as critical. Within the executive office of the presidency (EOP) is the White House Office (WHO). Within the WHO are those staff people with whom presidents work most closely and on whom they rely. These are usually people from a president's past, perhaps the campaign. Though the urge to draw exclusively from this pool of past confidants is understandable, presidents are wise to include in this circle those with experience in Washington, D.C., especially experience dealing with Congress. When presidents are elected as outsiders (Jimmy Carter in 1976, Bill Clinton in 1992), their staffing choices often include those without such experience. Carter's staff was informally called the Georgia Mafia, and Clinton's was Friends of Bill. Although this might increase presidents' comfort levels with the information and advice these staffers bring, they run the risk of making poor choices because their staffs may lack crucial expertise in understanding the unique context of the nation's capital.

The White House staffing structure as we know it came about with the advent of the modern presidency in the administration of Franklin Roosevelt, through a combination of congressional legislation authorizing a larger permanent staff and Roosevelt's executive order reorganizing the executive branch, creating the EOP and moving the Bureau of the Budget into the EOP. More is said later about the creation and use of executive branch agencies to grow presidential power.

The role of White House staff is to help the president make crucial day-to-day decisions in governing. As such, one of the most important functions they provide is to make sure information reaches the president in a timely fashion, in a method that suits their boss's particular working style and personality and that is carefully honed so as not to overwhelm him with an overabundance of superfluous detail. As the discussion of presidential personality illustrates, the work habits and preferences of presidents as human beings have an effect on their work as presidents and on the structure of the work environment in which they flourish.

White House staffs are organizations and as such can be structured in a variety of ways. They can be structured loosely without formal hurdles to presidential access. This style is often called spokes of the wheel or collegial; aides know they have access to the president and can freely give their advice and perspective. Although this puts a fair amount of pressure on presidents' time and resources, it ensures they are informed by a wide range of views and opinions and increases the volume of the data and information they personally see (Johnson, 1974).

A variant on the collegial spokes of the wheel is the competitive model. Here, staff members are given overlapping responsibilities, with the full knowledge that someone else is working toward a solution to the same problem as they are. This fosters competition, and the assumption is that the competition will produce the best effort. Franklin Roosevelt used this system.

By contrast, the office can be structured hierarchically, with a clear chain of command. In this configuration, decisions are taken by junior staff when those staffers have authority and resources to make the decisions. Only the difficult or controversial choices get pushed up the chain. There is a gatekeeper, usually called a chief of staff, who closely monitors the president's time and schedule. This ensures that the president is only involved in the most important issues and that smaller, more trivial matters are kept off his or her desk. The danger is that the filtering system will become too rigid and that important details or issues will be kept from the president. This can lead to a potentially damaging situation going unaddressed.

Alexander George (1980) argued for what he called multiple advocacy or a hybrid approach where competition is managed so as to get the quality information without the potential for a corrosive work environment that can emerge from a situation where a staffer might feel his or her job is in jeopardy with each assignment. This system requires not a chief of staff screening access to the president, but rather an honest broker who can effectively synthesize the best and worst of each point of view. This honest broker reduces resource strain and overload on the president. It also is crucial that the president not tip his or her hand as to initial preferences since the proclivity to say "Yes, Mr. President" is hard even for top staff members to resist.

These preferences for staffing structures go hand and glove with presidential preferences for the nature of the

information presidents receive. Presidents vary in the format for information: oral briefings in meetings versus written policy memos. They also vary in the quantity of data they wish personally to see. Sometimes a president wants the essence of a problem condensed to a one-page summary (e.g., Reagan), and some prefer in-depth briefing books (e.g., Carter).

The type of staffing structure presidents choose is one of the few things about the decision-making process they can control. There are many facets of this context that constrain their choices: past promises, preferences of advocates for a particular position whose support they may need in the future, and the sheer volume of important issues. Given this difficult environment, it is critical that these processes fit the president as a person.

The President and the Institutions of Government

The President and Congress

In many ways, Congress is the institution with which the president's ability to work effectively will be tested. Although there are nonlegislative vehicles for achieving the president's policy agenda (which are discussed later), the most enduring are laws. Additionally, history tends to shine a more favorable light on presidents who are able to pass large proportions of their legislative agendas. Presidents' agendas are indications of their priorities and on what they are willing to stake their reputations (Light, 1980).

The crafting of an agenda is a function of many factors (Light, 1980). First, a president has to consider commitments and promises from the campaign and weigh the costs versus benefits of trying to fulfill each. Breaking a high-profile promise that was salient to an important component of his or her electoral coalition will have greater consequences than breaking a less visible promise or an issue on which the country is divided. Additionally, once in office, the president may find that fulfilling the promise would have unforeseen and unintended consequences, and he or she may choose to moderate his or her position.

The climate of the times also bears on the choice of items on which to focus. In crises (economic or military), items can rise to the top of the agenda even if they were not mentioned in the campaign. The tragic events of September 11, 2001, illustrate this point. This threat to the country's national security precipitated a historic overhaul of the executive branch (in creating the Department of Homeland Security from an amalgam of existing agencies) as well as the far-reaching Patriot Act Congress passed in 2001.

As political scientists have tried to understand agenda creation and fulfillment, one of the methodological issues that arise is how to measure the agenda. Paul Light (1980) made use of the executive branch process for tracking legislation as it moves through Congress. In most administrations,

the Office of Management and Budget (OMB) regularly examines enrolled bills (bills that have passed both the House of Representative and the Senate and are headed to the president's desk for signing or veto) and determines whether they are in accord with the president's stated preferences. This in accord (I/A) list then serves as a list of issue preferences.

However, not all preferences are also priorities. To determine the issue on which the president would likely invest his or her resources, Light (1980) cross-referenced the I/A list with the State of the Union addresses. The resulting list was taken as the president's agenda. He supplemented these data with interviews from White House staffers to get a sense of the way in which the agenda is crafted as well as the factors that influence agenda success.

Light's (1980) central thesis is that a president starts an administration with a set of resources. Presidents are the most rested they will be at any point in their administrations, and goodwill and public support from the campaigns create a honeymoon with voters and members of Congress alike. These resources, which Light calls capital, will decline precipitously in an inverse relationship to increasing effectiveness—essentially learning on the job. These two cycles, decreasing influence and increasing effectiveness, mean that in order for a president to implement an agenda, he or she must move quickly—hit the ground running.

With an understanding of this and of the partisan context in Congress and informed by their own goals (historical, policy, and reelection), wise presidents will make good choices about whether to attempt to change the status quo, how big and costly a program to propose, and whether to attempt legislation or work within their powers of executive orders. Light's analysis confirms that presidents are more likely to be successful when they attempt smaller, less radical policies and move early in their first terms.

Indeed, scholarship in the area of presidential success with Congress mirrors Light's analysis in terms of the tools and resources available to presidents. Among the most important resources is the amount of party support in Congress. Presidents are much more likely to be successful during unified party government or when, in times of divided party government, members of Congress from their own parties exhibit party loyalty and cohesiveness.

Public support for presidents and for their policies can be useful. Going public, or attempting to use the bully pulpit to persuade the public and thus put pressure on members of Congress, is one possible tool (Edwards, 1983). Presidents Reagan and Clinton were especially skilled communicators. The White House can use the media and staged photo opportunities to rally support while crafting a message that attempts to paint the opposition into a no-win corner.

Although presidents (and members of the media) may believe it is an effective way to change opinion, data suggest that the influence of this going-public strategy is at

the margins. Presidents generally are unable to change public opinion significantly (Edwards, 1983). Recent scholarship in this area has focused on the details of presidential–congressional bargaining over policy. This line of research is premised on the notion that the president and the majority of Congress each has an ideal preferred policy, perhaps arrayed on an ideological continuum. The proximity of these preference points starts as a function of shared ideology and partisanship. However, the ability of the president to move Congress toward his or her preferred outcome is where skill and the use of informal resources play a part.

Presidents have a variety of carrots and sticks with which they can lobby members of congress, and the range of these options is greater among their fellow partisans. For example, working with congressional party leaders, a recalcitrant member might be threatened with a committee assignment or help in the next campaign. Conversely, members can be offered a range of goodies for their support: a presidential visit to their district or campaign support, for example.

At their core, however, members of Congress are driven by reelection goals. They are not going to sponsor legislation or vote in committee or on the floor in a way that hurts their public standing and that weakens them electorally. This is one of the reasons for the importance of presidential popularity. Popular presidents, not only who have high approval ratings but also who received a high percentage of the vote in the member's district or senator's state, more easily persuade members of Congress.

Thus, presidents cannot be faulted for trying to rally the public to their side. Another category of resources in bargaining with Congress is the range of options associated with the veto. The use of this formal power to achieve a president's policy ends can be seen as a show of weakness. That is, if presidents lack the political capital to push their legislative agendas through Congress, they can create legislative roadblocks through the veto. Presidents such as Gerald Ford, whose partisan support was extremely low and whose public support deteriorated after he pardoned his predecessor, Richard Nixon, used the veto often (he issued 66 vetoes in just over 2 years). Worse than the use of the veto, however, is to have one's veto overridden by Congress. President Ford, for example, had 12 of his vetoes overridden.

By contrast, presidents can inform Congress of their preferences and attempt to move them in their preferred policy directions by threatening to veto legislation (Cameron, 2000). These threats can take a variety of forms and can come from many executive branch sources. For example, the threat can be relatively nonspecific, which leaves the president room to negotiate: "The president will not be supportive of legislation that contains . . ." Or the threat can be explicit: "I will veto this bill if it remains in this form." Although the explicit threat provides clear direction to Congress on presidential preferences, if the

bluff is called, and Congress does not change the legislation, the president must veto the legislation or lose his or her bargaining position for the long term. Similarly, a threat issued by a White House staffer or cabinet secretary or even an anonymous quote to the media ("Staffers report . . .") leaves the president some degree of deniability and room to maneuver.

The President and the Executive Branch

One might think that the president's relationship with the executive branch is inherently less likely to be contentious than his or her relationship with Congress. After all, the president has appointment power over a significant portion of the executive branch, and the agencies are intended to carry out his or her policy directives. On the other hand, the executive branch's sheer size, its storehouse of substantive knowledge, and its relationship with congressional appropriators and with the groups it regulates create a source of power that is independent from the president.

The executive branch is composed of the 15 executive branch agencies and numerous boards, agencies, and government corporations, whose staff are specialized experts in their fields. While about one third of each agency is political appointments (appointed by the president and often requiring Senate confirmation), the vast majority of executive branch employees are civil servants. That means they have earned their jobs through the meritocracy of the civil service exam and are not politically beholden to the president.

In addition to this independence from the president, members of the bureaucracy have a nexus of relationships called the iron triangle, from which the president is largely excluded. The points of the triangle are the agency, the congressional committees, and the groups the agency regulates. One leg of this triangle is the regular exchanges an agency has with congressional committees who write legislation affecting how their jobs are performed. These same legislators are often the appropriators of the agencies' funds. For example, the day-to-day operations of the U.S. Department of Agriculture (USDA) can be affected by new laws dictating the business of ranchers and farmers. This legislation is crafted significantly in the House and Senate agriculture committees, which may convene hearings. USDA employees would likely be called as expert witnesses to these hearings owing to their substantive knowledge. Similarly, the USDA (as would all executive branch agencies) will take an active interest in the annual congressional budgeting and appropriating process, further cementing the relationship.

Another leg of the triangle is the relationship between the agency and the business, group, or industry that it regulates. In addition to individual members of these groups, in the case of the USDA, the ranchers and farmers, the agency regularly interacts with interest groups that lobby on behalf of these client groups. These regular

interactions further solidify the bureaucracy's independence from the president.

The President and the Courts

As with the other two branches of government, the president's relationship with the federal courts has both a formal and constitutional component and also a more informal or nuanced one. Among the formalized roles is the process of judicial selection. The president nominates federal judges, including Supreme Court justices, and the Senate has confirmation power. As with passing legislation, the dynamics of this exchange vary significantly depending on unified or divided party government. The mechanics of the process are also different for the lower federal courts (district courts and courts of appeals) than for the U.S. Supreme Court.

Senators have greater influence over the selection of lower federal court judges. The home senator—that is, the senators whose state is within the jurisdiction of the court to which the nominee would be appointed—would certainly be consulted by the White House in their selection process. Additionally, if a home senator opposes the appointment, he or she can effectively block it since other senators are likely to take their cues from him or her. Additionally, if the president's partisan opposition is in the congressional minority, they can use the filibuster to block nominees.

Blocking nominees, whether at the lower federal court level or to the U.S. Supreme Court, is most achievable for senators if the nominee is ideologically polarizing or has some weakness in his or her portfolio or if the president's opposition controls the Senate. However, White House staffs have become increasingly sophisticated in their selection process, especially since the Reagan administration. This decreases the likelihood of making a poor choice, someone who is not confirmable.

President Reagan established the standard that presidents since have used for vetting potential judicial nominees. The first stage in this process is to establish the potential nominee's intellectual and judicial credibility. Variables such as the law school attended and honors earned (such as editor of *Law Review*), clerkships, and other judicial appointments will all be considered. Because of the potential for opposition on partisan or ideological grounds, a nominee must first reach a threshold level of competence. In addition to this résumé fact-finding, a full investigation into the candidate's personal life will commence. Poor judgment, unethical or criminal behavior, or moral turpitude can all be fodder for an opposition's effective attack. In addition to partisan and ideological compatibility, presidents might have demographic preferences in their choices (e.g., to choose a woman, minority, or relatively younger justice). The intensity of these decisions is highest when the search is for a Supreme Court nominee.

Although the influence of the president on the federal judiciary and especially on the Supreme Court is most clear in terms of appointments, he or she does have other paths to affect judicial decision making, again especially at the Supreme Court. The Department of Justice, specifically the attorney general (AG) and solicitor general (SG), works closely with the Supreme Court. When the United States is a party to a case argued before the Supreme Court, either the AG or the SG (or one of their deputies) serves as counsel for the United States. The relative frequency of cases involving the United States creates a situation in which the AG and SG are repeat players before the Court. The sheer frequency of these appearances provides a measure of comfort before the bench that can be a tactical advantage in the legal proceedings.

Additionally, the Office of the Solicitor General (OSG) serves as a gatekeeper for all cases in which the United States is a party that loses at lower court and wishes to file a writ of certiorari, or a petition to have its case heard before the Supreme Court. Because of the volume of cases appealing for certiorari, the culling of nonmeritorious cases involving the United States by the OSG is invaluable to the Court (Caplan, 1987).

Perhaps one of the clearest avenues of executive branch influence before the Court is in filing of *amicus curiae* or friend-of-the-court briefs. Although any person or group may file such a brief (an essay outlining the support for a particular party in the case or line of legal reasoning), the frequency with which the OSG files such briefs creates a familiarity with arguments likely to be persuasive. Indeed, the success rate of the OSG (the times its position was on the same side in which the Court ruled) is extremely high, upwards of 75%. The Court thinks so highly of the work done by the OSG that on occasion it will request that the OSG weigh in and file an *amicus brief*, sometimes even going as far as to ask the OSG to comment in oral arguments. No other *amicus* has rights to make oral arguments (Pacelle, 2003).

Policy Implications

Public policy is often broken into three substantive areas: social, economic, and foreign policy. Each has its own contexts and challenges for presidential leadership. The president is but one player in the policy-making arena and the federal government but one of the many avenues for policy making. This section of the chapter draws on the previous discussions of Congress and the executive branch and the president's relationship with the public in order to make sense of the potential for a president to achieve his policy goals.

Social Policy

Because of the significant heterogeneity of issues within this broad category, research on the president's role

in social policy tends to be case study based, making it more difficult to draw generalizable findings. It is difficult to isolate with certainty specific factors that lend to presidential leadership across all areas of social policy. However, much of the discussion of Congress and decision making applies here. For example, when a president wishes to make a bold change in policy through lawmaking, he or she needs partisan support within Congress and public support and must be savvy in marshalling his or her resources at all stages of the legislative process. Conversely, if a president has few congressional partisans and little political capital, he or she is wise to keep a program small, incremental, and, ideally, self-executable within the structure of the executive branch.

The president has a variety of resources at his or her disposal at the various stages of the policy-making process: agenda setting, policy formulation, policy adoption, policy implementation, and policy evaluation. To be effective, he or she needs to have a good sense of the degree of consensus in the public on the degree of the problem and the need for a solution (agenda setting). Presidents will often try to influence public opinion on this, though the effect of this is marginal. The effect of the president on the policy formulation through evaluation stages is largely a function of the previous discussions on the president's relationship between the bureaucracy and Congress.

Economic Policy

Economic policy differs from social policy in that it is more likely either to be cyclical (like the annual budget), episodic (like tax policy), or precipitated by economic crisis (the 2009 Troubled Asset Relief Plan). Presidents have some degree of influence with the budget, though that would depend significantly on congressional support or opposition.

The range of ways in which the president's staff can assist him or her on economic policy belies the piecemeal way in which economic policy is made. The Department of the Treasury houses the Internal Revenue Service, which is responsible for collecting federal taxes. As such, Treasury has good information about projections on the amount of revenue, which plays a part in the budgeting process. Several factors affect the degree to which budgeting and revenue are strictly tied. The country goes through periods of fiscal conservatism where the mood of the country leans toward budget-neutral spending programs (if spending is increased on Project A, costs are reduced on Project B) or toward deficit reduction. The latter is difficult since it requires either a reduction in programs and services or an increase in taxes, rarely politically popular. A change in tax law, of course, requires congressional action.

Since the president's budget is the starting point for congressional negotiations, this is a potential area of influence. Not only is the budget cyclical, but also its creation requires input from across the executive branch and

significant coordination within the EOP. The president's budget is transmitted to Congress on or about February 1 of each year, in anticipation of the new fiscal year that begins October 1. The previous summer, cabinet-level agencies begin formulating their requests that will be sent to the OMB, which coordinates the requests in anticipation of transmission to senior White House staff. Although the formulation of agency budget requests is generally the responsibility of middle-level staff within each agency, once the request is sent to OMB, subcabinet and cabinet secretaries will begin to lobby the White House. Each secretary is advocating for his or her agency, to acquire the greatest resources in a usually constrained fiscal environment.

These internal White House budget negotiations necessitate good information about existing and forecasted economic trends. In addition to the revenue projections from Treasury and the impact of each agency request on the budget from OMB, the Council of Economic Advisors provides its expertise. This small group (about 25 to 30 staff and three board members) is responsible for generating macroeconomic data on the state of the economy, both currently and projected over several fiscal years.

In late December and January, the president and his or her closest advisers assimilate these data with his or her policy priorities. The budget that is submitted in February not only reflects these priorities but also reflects the White House's sense of likely congressional support or opposition. The successfulness of the president's budget is also affected by unified or divided party control of Congress and by electoral pressures on members of Congress. Lobby groups for the client groups affected by the budget decisions are skillful in applying pressure to members of Congress, especially those who serve on congressional committees charged with budgeting and with appropriating the funds.

Many of the criteria on which the health of the economy is judged are largely out of the president's control. For example, during times of slow economic growth, lowering interest rates can have the effect of freeing capital, thus nudging spending. This in turn can stimulate production, which can help bolster individual workers' paychecks or job prospects. However, the government entity that most affects interest rates is the Federal Reserve Board (the Fed). The members of the Fed's board of governors and the chair of the Fed are all appointed by the president but must be confirmed by the Senate. Their term of service (14 years) combined with a tradition of independence from political pressure all decrease the president's ability to control or influence this aspect of the economy.

Unemployment levels, another measure of economic health, are also out of the president's direct control. Although he or she crafts policies to encourage job growth, clearly he or she cannot will companies to hire more workers. This is an offshoot of a growing economy; when demand for goods and services increases, companies will hire.

Foreign Policy

The context of foreign policy making is quite different from social or economic policy. The president must be sensitive to the foreign policy orientation of members of Congress as well as the mood of the public at large. For example, are those members of Congress in key positions to block either policy initiatives or appropriations predisposed toward internationalism or engagement with the world community, or is the trend toward isolationism? The post-September 11th world of global terrorism and instability has created another set of challenges for the president in creating global alliances and in rallying domestic support for his policies.

Many aspects of foreign policy decision making constrain presidents. For example, although the president has the ability to commit troops and in some ways to take unilateral action, his or her ability to exert influence over foreign leaders is even less than his or her ability to pressure members of Congress. Additionally, where in economic and social policy making the president can often be overwhelmed with information, he or she must question his or her foreign policy sources. Having sufficient and accurate data on which to make decisions is critical and difficult.

The presidents have several policy-making tools and staffing structures to assist them. The Department of State takes the lead in international diplomacy, which can involve treaties and international agreements. Formal constitutional actions such as treaties require Senate ratification. The Senate also must confirm all cabinet members, which includes the secretaries of defense and state, as well as ambassadors. War-making power is also shared. Although Congress is entrusted with the power to declare war, the president is commander in chief and is responsible for setting policy direction for any military conflict or deployment of troops.

In addition to those of state and defense, the president has other advisory structures to assist him. On the military front, the joint chiefs of staff work closely with the secretary of defense. These are the heads of each military branch, and the chair of the joint chiefs can be a key policy adviser. The president also receives regular, usually daily, intelligence briefings from the Central Intelligence Agency, the National Security Council (NSC), the Director of Homeland Security, and the Defense Intelligence Agency. Traditionally, the NSC, headed by the national security adviser, is charged with coordinating these sometimes disparate streams of information.

Future Directions

The study of the presidency is dynamic, changing to reflect the evolution of the office and institution and to incorporate new analysis techniques. For example, the presidency of George W. Bush saw a substantial increase

in the use of an informal tool in presidential congressional bargaining: the signing statement. Like the veto threat, this is a mechanism for the president to achieve policy goals. However, unlike the veto threat, the signing statement is unilateral. In the message issued when the bill is signed, the president can direct his or her bureaucracy not to implement certain sections of the bill he or she has just signed into law. Although this is controversial, only until a case is brought to and accepted for hearing by the federal courts can the constitutionality of its use be established.

Conclusion

The presidency is a unique institution in the U.S. system of government. A paradox, the president is at once both powerful and dependent on other branches and on public opinion for success. He or she speaks with a singular voice, yet his or her ability to move public opinion significantly is limited. These contradictions, however, are the threads into which the most fascinating tapestry of the executive branch is woven.

References and Further Readings

- Barber, J. D. (1972). *The presidential character: Predicting performance in the White House*. Englewood Cliffs, NJ: Prentice Hall.
- Bond, J., & Fleisher, R. (1990). *The president in the legislative arena*. Chicago: University of Chicago Press.
- Burke, J. P. (2000). *The institutional presidency: Organizing and managing the White House from FDR to Clinton* (2nd ed.). Baltimore: Johns Hopkins University Press.
- Cameron, C. M. (2000). *Veto bargaining: Presidents and the politics of negative power*. New York: Cambridge University Press.
- Caplan, L. (1987). *The tenth justice: The solicitor general and the rule of law*. New York: Knopf.
- Cohen, D. B. (2002). From the Fabulous Baker Boys to the master of disaster: The White House chief of staff in the Reagan and G.H.W. Bush administration. *Presidential Studies Quarterly*, 32, 463-483.
- Edwards, G. C. (1983). *The public presidency: The pursuit of popular support*. New York: St. Martin's.
- Edwards, G. C. (1990). *At the margins: Presidential leadership of Congress*. New Haven, CT: Yale University Press.
- Fisher, L. (1991). *Constitutional conflicts between Congress and the president*. Lawrence: University Press of Kansas.
- George, A. (1980). *Presidential decision making in foreign policy: The effective use of information and advice*. Boulder, CO: Westview Press.
- Grossman, M. B., & Kumar, M. J. (1981). *Portraying the president: The White House and the news media*. Baltimore: Johns Hopkins University Press.
- Johnson, R. T. (1974). *Managing the White House: An intimate study of the presidency*. New York: Harper & Row.
- Kernell, S., & Popkin, S. L. (1986). *Chief of staff*. Berkeley: University of California.

- Kumar, M. J., & Sullivan, T. (Eds.). (2003). *White House world: Transitions, organization and office operations*. College Station: Texas A&M University Press.
- Light, P. C. (1980). *The president's agenda*. Baltimore: Johns Hopkins University Press.
- McDonald, F. (1994). *The American presidency: An intellectual history*. Lawrence: University Press of Kansas.
- Moe, T. M. (1993). *Presidents, institutions and theory*. In G. C. Edwards, J. H. Kessel, & B. A. Rockman (Eds.), *Researching the presidency: Vital questions, new approaches* (pp. 337-385). Pittsburgh, PA: University of Pittsburgh Press.
- Neustadt, R. (1960). *Presidential power and the modern presidents*. New York: Free Press.
- Pacelle, R. (2003). *Between law and politics: The solicitor general and the structuring of race, gender, and reproductive rights policy*. College Station: Texas A&M University Press.
- Ponder, D. E. (2000). *Good advice: Information & policy making in the White House*. College Station: Texas A&M University Press.
- Roosevelt, T. (1913). *The autobiography of Theodore Roosevelt* (Centennial ed.). New York: Schribner.
- Rossiter, C. (1960). *The American presidency*. New York: Harcourt Brace.
- Shull, S. A. (1999). *American civil rights policy from Truman to Clinton: The role of presidential leadership*. New York: M. E. Sharpe.
- Skowronek, S. (1997). *Politics presidents make: Leadership from John Adams to Bill Clinton*. Cambridge, MA: Belknap Press.
- Sullivan, T. (1991). The bank account presidency: A new measure and evidence on the temporal path of presidential influence. *American Journal of Political Science*, 35, 686-723.
- Sullivan, T. (2004). *The nerve center: Lessons in governing from the White House chiefs of staff*. College Station: Texas A&M University Press.
- Taft, W. H. (1916). *Our chief magistrate and his powers*. New York: Columbia University Press.
- Walcott, C. E., & Hult, K. M. (1995). *Governing the White House: From Hoover through LBJ* (Studies in Government and Public Policy). Lawrence: University Press of Kansas.
- White House. (n.d.). *The presidents*. Available at <http://www.whitehouse.gov/about/presidents>
- Wilson, J. Q. (1989). *Bureaucracy: What government agencies do and why they do it*. New York: Basic Books.
- Winther, J. (2008). *Frost/Nixon: Complete interviews* [Motion picture]. United States: Liberation Entertainment.

85

AMERICAN JUDICIAL POLITICS

SHANNON ISHIYAMA SMITHEY

Westminster College

The field of judicial politics began when scholars began to doubt that the decisions of judges were driven solely, or primarily, by the law. The legal argument has traditionally maintained that judges are like technicians, applying the law to the facts, so that the decisions they make are not based on their preferences or their emotions but on an expert reading of the law. Judicial politics insists that this account is both incomplete and misleading. The primary goal in the field is to explain what factors influence the decisions made by judges, particularly those serving on the U.S. Supreme Court and other appellate tribunals. There are other topics raised in the field of judicial politics, including the interaction between courts and other parts of the political system and the ability of judges to effect change outside their courtrooms. Although these are important and worthy subjects for study, this chapter focuses on judicial decision making, since decision making is the area in which the field has made the most progress in developing explanatory theories.

Theory

The field of judicial politics currently finds itself with three major theories competing to explain judicial decision making: attitudinal, strategic, and institutional. The origins of the field lie in the first half of the 20th century in the

arguments of the legal realists. The realists argued that judges create rather than merely apply law, and they do so based on their personal conceptions of desirable policy (Clayton & Gillman, 1999; Segal & Spaeth, 1993). Political scientists took up the question of political attitudes' influence on judicial behavior when C. Herman Pritchett (1941) published his study of the U.S. Supreme Court during the Roosevelt administration. Pritchett rejected the idea that Supreme Court justices were objective legal experts, claiming instead that "it is the primary attitudes of the majority of the Court which becomes private law" (p. 890). He analyzed the dissenting votes cast by Supreme Court justices, finding consistent patterns in the combinations of justices who dissented together, a pattern that would not have been expected if legal factors were the primary cause of judicial decisions. "Where there were divisions of opinion," he wrote, "they appear to be for the most part explicable in terms of the opinion of the respective judges on public policy" (p. 898).

This legacy continues in the attitudinal model, which is put forth most emphatically by Jeffrey Segal and Harold Spaeth (1993) in their book *The Supreme Court and the Attitudinal Model*. The attitudinal model rejects the claim that judges follow the law rather than make it. Instead, the model contends the following:

The Supreme Court decides disputes in light of the facts of the case vis à vis the ideological attitudes and values of the justices.

Simply put, Rehnquist votes the way he does because he is extremely conservative; Marshall voted the way he did because he is extremely liberal. (p. 65)

The attitudinal model's success at predicting the votes cast by Supreme Court justices made it the dominant view in judicial politics through the mid-1990s.

The strategic model of judicial behavior has emerged to challenge the attitudinal perspective. This school of thought, with its roots in Murphy's (1964) *Elements of Judicial Strategy*, argues that judges prefer to vote their sincere policy preferences, but aspects of their environment modify their ability to do so. Judges cannot act on their attitudinal preferences without regard to context; they can achieve their policy goals only with the cooperation of colleagues. They must take institutional rules into account if they want to achieve their preferred policy outcomes. For example, a justice on the U.S. Supreme Court who wants to hear a case has to convince at least three others to grant certiorari. Once the case is docketed and argued, to control the majority opinion, a justice will need the cooperation of the senior justice in the majority, who assigns the opinion, and at least four other justices, who might demand concessions for their agreement. Rather than being able to act sincerely, justices must consider the likely reactions of others in order to plot a strategic course, at least some of the time. Justices may even vote against their most preferred course of action if the odds are against achieving sufficient cooperation from their colleagues. This sort of logic is offered to explain why justices are less likely to vote to grant certiorari when they are in the minority and therefore unlikely to get their way from a decision on the merits. Justices may even vote in ways that seem to contradict their preferences if doing so allows them to minimize the damage done to their ideal outcome. Thus, advocates of the strategic model contend that "we cannot understand the choices justices make—from the decision on certiorari through the choice of policy in the majority opinion—without taking into account the strategic nature of the decision-making context" (Epstein & Knight, 1998, p. 79; see also Maltzman, Spriggs, & Wahlbeck, 2000).

Because judicial decisions are not self-enforcing, judges have to consider more than just the preferences of their colleagues. Courts need cooperation from others if decisions are to be successfully implemented. Implementation may hinge on the compliance from other policymakers, such as police officers, lower court judges, and school boards. It can also depend on the voluntary choices of private citizens, who may or may not want to consume the decision—for example, women may have the right to an abortion, but unless they seek abortions and unless doctors are willing to perform abortions, that right remains unrealized. Negative reactions to court decisions from the public complicate the picture further (see Johnson & Canon, 1984; Rosenberg, 1991). Resistance to unpopular decisions may range from

mere evasion to outright retaliation. Congress can override statutory decisions, remove jurisdiction over particular issues of law, or even propose amendments to override unpopular constitutional interpretations. Decisions that alienate the public increase the likelihood that Congress may resort to such measures or that the executive may decline to enforce a decision. The strategic model argues that judges will take these external factors into account, anticipating possible reactions and acting in ways designed to minimize the possibility of backlash. Since judges must adapt their behavior to avoid external threats and to accomplish what is possible under the circumstances, the model rejects the idea that votes always reflect sincere preferences. It is this contention that brings the strategic model into direct conflict with the attitudinal model's emphasis on votes as indicators of judicial preferences.

Although the debate between the attitudinal model and the strategic model is the most prominent theoretical conflict in the field at present, there is a third model, usually dubbed *historic institutionalism*. While the strategic model and the attitudinal model focus on the policy goals of individual judges, the institutional model broadens its focus to the Court as a whole. Though admitting that individual concerns are important, the model sees institutional-level concerns (such as the power of the judiciary as a whole) as equally important for a full understanding of judicial politics. For example, to appreciate the role of the Supreme Court during the Civil War, it may be more important to understand the interaction between the Taney Court and the Lincoln administration than to assess the preferences of the individual Taney Court justices. Or one might follow Dahl's (1957) path, exploring the willingness and ability of the Supreme Court to resist the policies of other policy-making institutions. From the institutional perspective, judicial politics stands to gain more by "focusing less on the policy preferences of particular justices and more on the distinctive characteristics of the Court as an institution, its relationship to other institutions in the political system" (Clayton & Gillman, 1999, p. 3).

In addition to exploring interbranch relations over time, the institutional model documents changes in internal court practices and the difference these changes have made. The idea here is that when institutional norms and rules change, judges modify their behavior in response (and for more than just strategic reasons). For example, before John Marshall became chief justice it was common for multiple justices to write opinions in each case. Once Marshall introduced the practice of a single majority opinion for the Court, a consensus norm evolved that discouraged the practice of dissenting. Conversely, a series of changes in institutional context between 1925 and 1945 fostered an increase in the filing of concurrences and dissents. The Judges Bill of 1925 gave the Court discretionary control over its docket, increasing the percentage of controversial cases the Court would hear and providing more

time to consider them. At the same time, technological changes like the use of typewriters allowed for greater circulation of draft opinions. During the Stone Court, the justices began hiring an additional law clerk, and the Chief Justice was more tolerant of divergent opinions than his predecessors. All these changes encouraged the preparation and filing of more individual opinions (for further discussion, see O'Brien, 1999). The institutional approach gains the historical label in part from its emphasis on tracing such changes in institutional rules and the resulting adjustment in judicial behavior over time.

The institutional model emphasizes the effects of institutional context on judicial decisions, but in a somewhat different way than the strategic model. The argument is that institutional structures create strategic opportunities for justices seeking to advance their policy goals, but they do more than that. Institutions are made up of norms and rules that provide systems of meaning, including a sense of identity and responsibility for participants, and a logic of appropriateness that may foster concern for the institution and inhibit pursuit of personal goals. It may be helpful to think of the rules by which we play games. For example, the strategic model might stress the way that basketball players adapt to the rules in order to score as many points as possible (allowing them to increase their salary demands), while institutionalists point to passing, guarding, and other aspects of team play that benefit the group more than the individual. Institutionalists contend that judges develop normative attachments to their courts and act as stewards of their power. While advocates of the strategic model view judicial efforts to protect a court's "institutional authority . . . as a mean to an end—a policy end" (Epstein & Knight, 1998, pp. 48–49), the institutional model posits that judges are likely to be concerned with the welfare of the court for its own sake.

Clayton and Gillman (1999) argue the following:

Institutions not only structure one's ability to act on a set of beliefs; they are also a source of distinctive political purposes, goals and preferences. . . . With respect to Supreme Court politics, this means that the justices' behavior might be motivated not only by a calculation about prevailing opportunities and risks, but also by a sense of duty or obligation about their responsibilities to the law and the Constitution and by a commitment to act as judges rather than legislators or executives. (p. 5)

This reintroduces the idea that law may influence judicial decision making, since fidelity to law is an important part of the judicial role. This brings the institutional model into sharp conflict with the attitudinal model's insistence that it is not law but the political preferences of the judges that drive judicial decision making. Of course, the institutional model does not suggest that personal attitudes do not matter, but it does insist that "judicial values and attitudes are shaped by judges' distinct professional

roles, their sense of obligation, and salient institutional perspectives" (Clayton, 1999, p. 32).

Applications and Empirical Evidence

In developing and testing these theories, judicial scholars have followed a number of practices employed in Pritchett's (1941) analysis of the Roosevelt Court. The primary dependent variable has been the votes justices cast, and the main argument has been that differences in voting patterns reflect differences in judges' underlying political values or attitudes. Like many other political scientists at midcentury, judicial scholars adopted the tenets of behavioralism, which included the following beliefs: that political science should focus on phenomena that can be observed, that data should be quantifiable, that research should be directed by theory, and that prediction and explanation should be the focus of research (Segal & Spaeth, 1993). Votes could be counted, and the division among judges' votes could be compared to test the theory that attitudes drive those divisions.

Pritchett's (1941) study also excluded cases in which the Supreme Court had been unanimous. Pritchett suggests that different factors affect the justices' votes in unanimous cases:

In the great majority of these ballots the decision is unanimous. In such cases, presumably, the facts and the law are so clear that no opportunity is allowed for the autobiographies of the justices to lead them to opposing conclusions. (p. 890)

It seems a bit ironic that the foundational article of the attitudinal model would argue that such a significant percentage of judicial decisions were based on legal factors. One might have argued instead that unanimous cases reflected a high degree of attitudinal convergence among the justices; there is no clear way to establish one explanation as better than the other. What is clear is that a case in which every justice votes the same way provides no within-case variation to explain. Given the difficulty of knowing how to categorize unanimous cases, it quickly became standard practice to exclude them when studying attitudinal differences among the justices.

Some of the earliest behavioral work in judicial politics linked differences in voting to differences in judges' social background factors. The logic of this approach was that judges from similar backgrounds would have had similar life experiences, giving rise to different political attitudes and different voting behavior from judges with different backgrounds. Typical background factors include religion, social class, status of law school attended, and political party. Judicial scholars have largely abandoned the study of background factors since Ulmer (1986) demonstrated their limited utility. However, increasing diversity in federal judicial appointments has led to exploration of whether

significant differences exist among the decisions of judges of different races or between the decisions of male and female judges.

Scholars also documented the degree to which Supreme Court justices voted in consistent patterns, in blocs or in ideological alignment (for example, see Schubert, 1965). A bloc occurs when two or more justices vote together at a rate at least one standard deviation above the mean for the Court as a whole, and it is taken to indicate a high degree of attitude sharing among bloc members. Ideological alignments were usually established using Guttman scaling, a technique that assesses consistency in the ideological position of the justices relative to one another. To refine the analysis, judicial scholars often broke cases down by issue type to assess the degree of attitudinal overlap between justices in specific policy areas, such as civil liberties, federalism, and economic regulation. Currently, attitudinal studies tend to array the justices along a single underlying ideological dimension, from liberal to conservative, using the relative position of the justices of the Supreme Court to predict their votes.

The attitudinal model has had significant success in demonstrating a link between the ideological position of the justices and their votes, particularly in civil liberties cases. Segal and Spaeth (1993) found a correlation of .71 (adjusted R^2 of .60) between the ideological values of justices and their votes in the civil liberties cases formally decided by the Supreme Court between 1953 and 1989. Attitudinal studies have also had significant success in modeling decisions in specific areas of law where there have been enough cases to allow for statistical analysis. By exploring the interaction of attitudes and case facts, Segal and Spaeth successfully predicted 74% of justices' votes in search and seizure cases decided between 1962 and 1989, a significant improvement over prediction based on case facts alone.

The null hypothesis in these studies was the claim traditionally made by judges and lawyers—that judicial decisions were driven by the law. The legal model of judicial decision making includes a variety of possible legal determinants of judicial decisions—the plain meaning of the words in a law, the intentions of those who adopted it, or the legal doctrines established by earlier court decisions in the same area. Judicial politics scholars have investigated each of these legal rationales advanced by judges in their opinions and found them wanting. Legal models underestimate the degree of ambiguity in the language of statutes and constitutional provisions, as well as the degree of difficulty in determining the intention behind their adoption. The records needed to divine legislative or framers' intent may well be lost to history. Even if one has such records, neither legislatures nor constitutional conventions are unitary actors with a single intent. Thus, the search for intent is often futile, and judicial scholars find it difficult to believe that such ambiguous guidance can really determine case outcomes.

Even when legal language seems clear, judges do not always follow it. The justices of the Supreme Court have

constructed limits on rights like freedom of speech, which appear absolute in the text, while finding rights, like the right to privacy and travel, which are not mentioned in the text at all.

If the Court can read rights out of the Constitution that it explicitly contains while simultaneously reading into the Constitution rights that it does not explicitly embrace, the plain meaning rule obviously fails as an explanation of what the Court has done. (Segal & Spaeth, 1993, p. 38)

The attitudinal evidence also indicts the idea that either framers' intent or precedent determines case outcomes. Arguments based on framers' intent and precedent are offered on both sides of most cases before the Court. In fact, the majority and dissent often refer to the very same precedent while reaching exactly opposite results. If judges on both sides of a case claim to be governed by the same legal factors at the same time, despite reaching diametrically opposed conclusions, law appears to be a rationalization rather than a causal factor (Segal & Spaeth, 1993).

A modified version of the legal model includes judicial role conceptions: the idea that judges follow the law because they believe that is what their job demands of them. The role conception explanation leaves open the possibility for attitudinal influences, since judges vary in the degree to which they believe it appropriate to act on their own preferences. Judges are sometimes explicit in contrasting what they would like to do with what they feel the judicial role allows them to do. For example, Justice Harry Blackmun voted to sustain constitutionality of the death penalty despite his declaration: "I yield to no one in the depth of my distaste, antipathy, and indeed abhorrence, for the death penalty. . . . Were I a legislator, I would vote against the death penalty" (*Furman v. Georgia*, 1972, pp. 405–406). The greatest limitation on judicial attitudes ought to come from the restraintist role, since that perspective views any personal or political influence on judges' decisions as illegitimate. Gibson (1978) did find significant differences in sentencing behavior based on role conceptions among trial court judges. His analysis revealed significant correlations between the political views of activist judges and their sentencing decisions but no significant correlation between political views and sentencing decisions among those advocating restraint. Attitudinalists like Segal and Spaeth (1993), however, tend to dismiss the idea that appellate judges are restrained by role conceptions. Their argument is that justices often claim to be restrained when they are really voting to uphold policies they like. Judicial scholars in the attitudinal tradition therefore discount the argument that law plays a determinative role in judicial decisions.

Despite the predictive success of the attitudinal model and the absence of empirical tests supporting the legal model, not everyone is convinced that judicial decisions are explained by politics instead of law. Certainly judges continue to claim that their decisions are driven by legal

texts, framers' intent, and precedent. Law review articles base their explanations of court decisions on these same factors, with particular emphasis on precedent. Law school curricula stress the study of precedent and how to apply it (Clayton, 1999; Segal & Spaeth, 1999). Even some political scientists argue that precedent affects certain aspects of judicial decisions. For example, the strategic model stresses that the use of precedent has strategic benefits for judges. For example, Epstein and Knight (1998) contend that justices of the Supreme Court follow precedent because it increases the odds that future justices will respect the precedent made by the current justices and because society grants legitimacy to judicial power when people perceive the Court to be conforming its decisions to existing case law. Judges may rely on precedent because they believe it increases the odds that others will perceive their decisions as legitimate and comply with them.

Segal and Spaeth (1999), however, remain doubtful, finding few instances in which Supreme Court justices shift their positions in an area of law to conform to precedents with which they initially disagreed. Of course, Segal and Spaeth do not count decisions in which justices continue to follow precedents decided before they ascended the bench nor those with which they initially agreed, since in such cases there is no clear way to establish that precedent is restraining a contrary personal preference. This move, although understandable on behaviorist grounds, eliminates from the test of precedent's influence the vast majority of instances in which the justices act in ways that are consistent with precedent. Institutionalists are more likely to conclude that precedent has some influence on the development of judicial decisions. Institutionalists argue that while judges are influenced by their policy preferences, they are also qualitatively different from other policymakers. Judges are socialized to revere precedent. The arguments they hear from lawyers and other judges are freighted with precedent, and judges are expected to conform their decisions to its dictates. Precedent may not determine the vote, but it may nevertheless influence case outcomes by affecting the array of options presented to judges and helping them distinguish between legitimate and illegitimate policy alternatives.

Based in part on the idea that the acid test of a scientific model is its predictive abilities, a group of researchers recently conducted a head-to-head test to compare the ability of legal experts and political science models to predict the outcomes of Supreme Court cases during the 2002 term ("Symposium," 2004). The political model used decision trees based on a variety of political factors, including ideological direction of the lower court ruling, type of parties to the case, the legal issue raised, and the likely positions of other justices. The legal experts were law professors or highly specialized attorneys, several of whom had clerked for the Supreme Court, who were asked to predict case outcomes in up to three cases in their areas of expertise (for further details, see Washington University Law, 2002).

The political science model was more successful in predicting case outcomes (75% to 59.1%), while the legal experts were slightly more successful in predicting the votes of individual justices (67.9% to 66.7%). The legal experts did particularly well relative to the political model in technical areas of the law such as administrative procedure. From this, Epstein ("Symposium," 2004) concludes that each perspective has something to learn from the other: "What the results from the project seem to suggest is that legal academics who stress principles and neglect politics in their explanations of case outcomes do so at their own peril," but the success of legal academics in predicting individual justices' decisions "suggests that political scientists concerned with explaining the range of judicial decisions can no more afford to neglect law than law professors can ignore politics" (p. 757). Thus, scholars have found that there is something to be gained from combining approaches to gain a more complete understanding of judicial behavior.

Something may also be gained from combining aspects of the three main models of decision making. For example, when it comes to understanding the assignment of majority opinions on the Supreme Court, the effects of ideology, institutional norms, and strategic calculation are all apparent. Opinion assigners are more likely to assign important cases to ideologically similar colleagues, but they must also take into account the norm of relative parity in workload among the justices (Maltzman et al., 2000; Segal & Spaeth, 1993, chap. 6). When the Court is closely divided, the assigner may also wish to assign to the marginal member of the majority coalition to shore up the likelihood of having the case decided in the preferred direction. Emphasizing only one theory to describe such behavior leaves too much unexplained.

Similarly, strategic and institutional approaches have offered convincing accounts of individual cases in which Supreme Court justices sacrificed their personal preferences in the face of threats to the institution. *Marbury v. Madison* (1803), *Ex Parte McCordle* (1867), and *Barenblatt v. United States* (1959) all provide examples of courts backing down in the face of a hostile and threatening Congress (Epstein & Knight, 1998, chap. 5). Though they accept the separation of powers story in a limited number of cases, attitudinalists emphatically reject the idea that the Supreme Court regularly takes the preferences of other branches into account in the typical case. Although there is some evidence that the Court responds to the influence of the solicitor general, in the aggregate attitudinal evidence seems to explain much of Supreme Court voting quite well. In head-to-head tests, the attitudinal model outperforms models based on strategic calculation related to the separation of powers. For example, using a pooled cross-sectional time series analysis for 1947 through 1992, Segal (1997) finds that ideology is consistently better at predicting justices' votes than a variety of separation of powers models. But the ability of the justices to vote their preferences depends on the Court's

institutional position. In fact, as Segal and Spaeth (1993) admit, the relative freedom of the justices to ignore strategic considerations may itself be a function of the strategic position in which the modern Supreme Court finds itself. The Court has not always enjoyed its current level of insulation from external political pressures. Strategic and institutional models can contribute an appreciation for the way the Court's strategic position has changed over time, including strategic retreats as well as current strengths as part of a larger story of institutional development. This kind of analysis may also facilitate comparative work, since variation in the protections provided to judicial independence in other countries could be compared with that in the United States.

It may also be worthwhile to combine the three models in exploring the behavior of judges on lower courts. Although there may be few external constraints on the ability of Supreme Court justices to act on their preferences, the judges on the federal courts of appeal may be more constrained. Their institutional position below the Supreme Court means that appeals court judges can be overturned on appeal. This may explain why they evince relatively high rates of adherence to precedent (Songer, Segal, & Cameron, 1994). Strategic calculations may also influence the way appeals court judges interact with their colleagues, though currently ideological distance is seen as a better predictor of an appellate judge's decision to dissent (Hettinger, Lindquist, & Martinek, 2004).

Differences in the influence of public opinion on judicial decision making are also apparent depending on the level of court examined. Despite its reputation for protecting minorities against the majority, the Supreme Court has been shown to reach decisions that match reasonably closely with public opinion (Marshall, 1989; Mishler & Sheehan, 1993). This correlation between the Court's decisions and public opinion does not prove causation; although it is possible that the justices respond directly to public opinion, it may well be that the justices merely share the views of their contemporaries among the public or that public opinion filters through to the Court as a result of the appointment process (Dahl, 1957; Norpoth & Segal, 1994). No one has yet been able to establish beyond doubt that public opinion actually drives case outcomes at the Supreme Court level.

The data are more supportive of a causal relationship among state judges. The public has no direct role to play in electing or retaining federal judges, but voters are involved in selection and retention of most state judges. Voters tend not to be too knowledgeable about candidates in judicial elections, and most judges are reelected by significant margins (Aspin & Hall, 1994; Hall, 2001). Even so, scholars have been relatively successful in demonstrating a link between public opinion and judicial behavior at the state and local levels. For example, Gibson (1980) demonstrated that Iowa judges who rode circuit decided cases differently in different courthouses and that these differences conformed to differing community norms. Kuklinski

and Stanga (1979) found that length of sentence judges handed down in marijuana cases moved in the direction of their county's election results after a referendum was held on decriminalization. Such behavior might be explained by a democratic role conception among judges, but scholars are more likely to attribute such shifts to strategic calculation. Judges want to keep their jobs, so even the possibility of electoral defeat can motivate them to stay close to public opinion (Hall, 1992).

Just as the same factors may not have the same effects in all courts, the same factors may also work differently depending on changes in context. There is an encouraging trend in recent research to emphasize how changes in context can change the impact of causal factors on judicial behavior. More nuanced models have demonstrated that the impact of electoral pressures on judges need not be the same across an entire judicial career but that judges are more likely to change their sentencing decisions when cases are closer to the next election (Huber & Gordon, 2004). Rather than exploring the connection between public opinion and outcomes in all cases, McGuire and Stimson (2004) distinguish the effects of public opinion in cases based on whether the Court is overturning or affirming rulings of the lower courts. More precise specification of this sort is most welcome.

Policy Implications

That political factors influence judges' decisions seems certain, even if there continues to be disagreement as to the exact extent of that influence. The next question is what should be done with that information. One area where this knowledge might be put to good use is the process of evaluating appointments to the Supreme Court and the federal courts of appeal. Clearly politicians are aware of the influence that judges' political views are likely to have on their decisions. Presidents have a tendency to select like-minded nominees whom they consider prone to decide cases in ways that will please them and their supporters, and senators have a tendency to base their confirmation votes on the same criteria (Epstein & Segal, 2005; Segal & Spaeth, 1993, chap. 4). At the same time, it is usual for presidents and senators to claim that nominees are chosen and evaluated on purely legal qualifications.

It would be refreshing if the findings about the link between attitudes and decisions could inspire greater candor from all parties to the nomination process. Judicial nominees could provide much more insight into the kind of decisions they would produce if they were more forthcoming during their confirmation hearings. Judicial reluctance to discuss hypothetical cases, combined with the failure of Robert Bork's nomination after he openly discussed his views at his 1987 confirmation hearings, however, suggest that we are unlikely to experience this sort of confirmation-hearing conversation any time soon.

Future Directions

Research into judicial decision making has benefitted from the availability of two major sources of data about the decisions of the Supreme Court and the federal courts of appeal. The ability to develop models predicting Supreme Court voting behavior was facilitated by the creation of the Supreme Court Database, which provides data classifying every vote cast by each Supreme Court justice in every case argued between the 1953 and 2008 terms. The database was created by Harold Spaeth with a grant from the National Science Foundation and made available for general use in the late 1980s. The current database houses 247 pieces of information for each case, including information on the path cases take to get to the Court, legal issues presented, the direction of justices' votes, authorship of opinions, and case outcomes (for further information, see the Supreme Court Database website). This greatly expanded the ease with which scholars can produce studies of the Supreme Court.

Similarly, the Judicial Research Initiative has significantly eased research into the decisions of the Federal Courts of Appeal. The Appeals Court Database provides information for a randomly selected sample of decisions reached by the federal courts of appeal between 1925 and 2002. There is also a Phase II data set that includes every court of appeal case that was subsequently reviewed by the Supreme Court, allowing for comparison between the decision of the Supreme Court and the lower court. Donald Songer was the original primary investigator for the Appeal Court Database. Ashlyn Kuersten and Susan Haire extended the dataset, and Ashlyn Kuersten and Todd Curry are responsible for flipping the data so that it is available by individual judge's vote as well as by case (for further information, see the Judicial Research Initiative website). Both of these provide a wealth of opportunities for scholars.

One issue that needs further attention from judicial scholars is the question of how to measure the concept of attitudes. Initially, attitudes were inferred from voting alignments without any independent measure of attitude itself. To improve the measurement of attitudes, the field has adopted Segal and Cover scores, which use newspaper editorials to generate a measure of each new Supreme Court justice's ideological position. Since this score is a snapshot of the justice's ideology at the time of nomination, it is an inexact measure of judicial ideology over time. Since there is considerable evidence that justices' attitudes are not static over time (Epstein, Hoekstra, Segal, & Spaeth, 1998), this means the field has significant specification problems in regard to the primary independent variable. One can of course measure ideology by using the percentage of the time a judge votes liberally or conservatively, but that leaves scholars inferring attitudes from votes again. The problem is compounded by the fact that the ideological position of the cases changes as well (Ulmer, 1981). The field has made progress in dealing with this issue. For example, cases are now often Baum corrected, referring to

Baum's (1992) practice of comparing the degree to which justices vote liberally in one period to the degree to which they voted liberally in earlier years (also see Epstein et al., 1998; Martin & Quinn, 2002). Nevertheless, more progress needs to be made in assessing the ideological stimulus points of the cases and the ideological positions of the justices relative to them.

Another issue that needs to be addressed is the narrowness of the main dependent variable. Votes clearly matter, but so do the opinions that judges write to explain those votes. Cases present more than yes-or-no opportunities for policy making. Policy options emerge from a variety of sources: legal briefs, lawyers' arguments, and justices' own (or their clerks') legal research. Opinions do more than provide rationalization for policy outcomes. Opinions provide insight into the way judges parse their options—opinion writers will offer some doctrines as justification for their decision, while rejecting others. The choices contained within these opinions matter. They provide guidance to attorneys and lower court judges about the rules to be used to resolve other cases. Opinions also signal future litigants about the direction of judicial preferences, signals to which interest groups have been seen to respond in order to advance their own political agendas through the courts (Behuniak, 1991; Kobylka, 1987). Although votes are important, they are far from the only important output of judicial decision making. More research needs to be done on the factors that influence judges in sorting through competing policy choices and crafting opinions.

Judicial politics needs to broaden its approach to the issue of judicial motivation as well. The attitudinal and strategic models (which encompass the majority of scholars in the field) both visualize judges as motivated exclusively by policy goals. As Baum (2006) points out, most explanations of judicial behavior “share a major limitation. Each model portrays Supreme Court justices (and, in some formulations, judges on other courts) as single-minded seekers of good legal policy, whether that means good policy or some combination of good law and good policy” (p. 14). If this were true, Baum says, jurists would be fundamentally different from other human beings, whose actions are affected by a variety of motivations. Most human beings are influenced by a combination of concerns—for career aspirations, financial gain, and workload control, among others. Baum concludes that one important part of judicial motivation is the desire to “maintain the esteem of people . . . who are important to them. They want not just to make good legal policy but to be perceived as doing so” (p. 159). This means that judicial actions should be interpreted as reflections of their sensitivities to reactions from a variety of audiences, including colleagues, social and professional groups, journalists, and legal academics. Some iterations of the institutional model go beyond the traditional emphasis on policy goals, including motivations based on institutional stewardship. Baum is nevertheless on firm ground in arguing that the dominant concept of motivation in judicial politics is overly narrow.

References and Further Readings

- Aspin, L. T., & Hall, W. K. (1994). Retention elections and judicial behavior. *Judicature*, 77, 306–315.
- Baum, L. (1992). Membership change and collective voting change in the U.S. Supreme Court. *Journal of Politics*, 54, 3–24.
- Baum, L. (2006). *Judges and their audiences: A perspective on judicial behavior*. Princeton, NJ: Princeton University Press.
- Behuniak, S. M. (1991). Friendly fire: Amici curiae and *Webster v. Reproductive Health Services*. *Judicature*, 74, 261–270.
- Clayton, C. W. (1999). The Supreme Court and political jurisprudence: New and old institutionalists. In C. W. Clayton & H. Gillman (Eds.), *Supreme Court decision making: New institutionalist approaches* (pp. 15–41). Chicago: University of Chicago Press.
- Clayton, C. W., & Gillman, H. (1999). *Supreme Court decision making: New institutionalist approaches*. Chicago: University of Chicago Press.
- Dahl, R. A. (1957). Decision making in a democracy: The Supreme Court as a national policy maker. *Journal of Public Law*, 6, 279–295.
- Epstein, L., Hoekstra, V. J., Segal, J. A., & Spaeth, H. J. (1998). Do political preferences change? A longitudinal model of U.S. Supreme Court justices. *Journal of Politics*, 60, 801–818.
- Epstein, L., & Knight, J. (1998). *The choices justices make*. Washington, DC: CQ Press.
- Epstein, L., & Segal, J. A. (2005). *Advice and consent: The politics of judicial appointments*. New York: Oxford University Press.
- Epstein, L., Segal, J. A., & Spaeth, H. J. (2001). The norm of consensus on the U.S. Supreme Court. *American Journal of Political Science*, 45, 362–377.
- Franklin, C. H., & Kosaki, L. C. (1989). The republican school master: The Supreme Court, public opinion, and abortion. *American Political Science Review*, 83, 751–771.
- Furman v. Georgia, 408 U.S. 238 (1972).
- Gibson, J. L. (1978). Judges' role orientations, attitudes, and decisions: An interactive model. *American Political Science Review*, 72, 911–924.
- Gibson, J. L. (1980). Environmental constraints on the behavior of judges: A representational model of judicial decision making. *Law and Society Review*, 12, 343–370.
- Hall, M. G. (1992). Electoral politics and strategic voting in state supreme courts. *Journal of Politics*, 54, 427–446.
- Hall, M. G. (2001). State supreme courts in American democracy: Probing the myths of judicial reform. *American Political Science Review*, 95, 315–330.
- Hettinger, V. A., Lindquist, S. A., & Martinek, W. L. (2004). Comparing attitudinal and strategic accounts of dissenting behavior on the U.S. Courts of Appeal. *American Journal of Political Science*, 48, 123–137.
- Huber, G. A., & Gordon, S. C. (2004). Accountability and coercion: Is justice blind when it runs for office? *American Journal of Political Science*, 48, 247–263.
- Johnson, C. A., & Canon, B. C. (1984). *Judicial policies: Implementation and impact*. Washington, DC: CQ Press.
- Judicial Research Initiative: <http://www.cas.sc.edu/poli/juri/JuRI>
- Kobylka, J. F. (1987). A court created contest for group litigation: Libertarian groups and obscenity. *Journal of Politics*, 49, 1161–1178.
- Kuklinski, J., & Stanga, J. (1979). Political participation and government responsiveness. *American Political Science Review*, 73, 1090–1099.
- Maltzman, F. J., Spriggs, F., II, & Wahlbeck, P. J. (2000). *Crafting law on the Supreme Court: The collegial game*. New York: Cambridge University Press.
- Marshall, T. (1989). *Public opinion and the Supreme Court*. Boston: Unwin Hyman.
- Martin, A. D., & Quinn, K. M. (2002). Dynamic ideal point estimation via Markov Chain Monte Carlo for the U.S. Supreme Court. *Political Analysis*, 10, 134–153.
- McGuire, K. T., & Stimson, J. A. (2004). The least dangerous branch revisited: New evidence on Supreme Court responsiveness to public preferences. *Journal of Politics*, 66, 1018–1035.
- Mishler, W., & Sheehan, R. S. (1993). The Supreme Court as counter majoritarian institution? The impact of public opinion on Supreme Court decisions. *American Political Science Review*, 87, 87–101.
- Murphy, W. (1964). *Elements of judicial strategy*. Chicago: University of Chicago Press.
- Norpoth, H., & Segal, J. A. (1994). Comment: Popular influence on Supreme Court decisions. *American Political Science Review*, 88, 711–716.
- O'Brien, D. M. (1999). Institutional norms and Supreme Court opinions: On reconsidering the rise of individual opinions. In C. W. Clayton & H. Gillman (Eds.), *Supreme Court decision making: New institutionalist approaches* (pp. 91–113). Chicago: University of Chicago Press.
- Pritchett, C. H. (1941). Divisions of opinion among justices of the U.S. Supreme Court, 1939–1941. *American Political Science Review*, 35, 890–898.
- Rosenberg, G. (1991). *The hollow hope: Can courts bring about social change?* Chicago: University of Chicago Press.
- Schubert, G. (1965). *The judicial mind*. Evanston, IL: Northwestern University Press.
- Segal, J. A. (1997). Separation of powers games in the positive theory of Congress and courts. *American Political Science Review*, 91, 28–44.
- Segal, J. A., & Spaeth, H. J. (1993). *The Supreme Court and the attitudinal model*. New York: Cambridge University Press.
- Segal, J. A., & Spaeth, H. J. (1999). *Majority rule or minority will*. New York: Cambridge University Press.
- Slotnick, E. E., & Segal, J. A. (1998). *Television news and the Supreme Court: All the news that's fit to air?* New York: Cambridge University Press.
- Songer, D. R., Segal, J. A., & Cameron, C. M. (1994). The hierarchy of justice: Testing a principal agent model of Supreme Court Circuit Court interactions. *American Journal of Political Science*, 38, 673–696.
- Supreme Court Database: <http://scdb.wustl.edu>
- Symposium: The Supreme Court forecasting project. (2004). *Perspectives on Politics*, 2, 757–793.
- Ulmer, S. S. (1981). Parabolic support of civil liberties claims: The case of William O. Douglas. *Journal of Politics*, 41, 634–639.
- Ulmer, S. S. (1986). Are social background models time bound? *American Political Science Review*, 80, 957–967.
- Washington University Law. (2002). *Supreme Court forecasting project 2002: Project description*. Available at <http://wusct.wustl.edu>

86

AMERICAN BUREAUCRACY

WILLIAM J. MILLER

Southeast Missouri State University

JULIA BECKETT

The University of Akron

This chapter summarizes the essential aspects of American bureaucracy for political science students during the 21st century. The four parts of this chapter are as follows: the theoretical bases, applications and empirical evidence, policy implications of bureaucracy, and continuing issues into the future as they relate to American bureaucracy. American bureaucracy refers to the administrative apparatus that accomplishes the tasks of government not otherwise conducted by the legislative or judicial branches. Other terms for bureaucracy include the administrative state, public administration, or public management. By examining the positive aspects of a well-functioning bureaucracy and acknowledging the myriad of critiques, this chapter demonstrates the important role of American bureaucracy in governance and policy. The inherent tension regarding bureaucratic performance that appears repeatedly in the history of American bureaucracy is that when functioning properly, bureaucracy is designed to operate without being noticed.

Bureaucratic and Governmental Theory

What Is Bureaucracy?

When political scientists speak of bureaucracy, they typically mean two things. First, they acknowledge the depth and breadth of bureaucracy, much as the initial definition

presented. Bureaucracy here is seen as an administrative apparatus that fulfills the duties of the executive branch of a government. It is the accomplishing arm of government. The second approach to bureaucracy taken by political science is a narrower, organizational definition that is often linked to a particular agency and is associated with Max Weber (1978).

Weber (1978) described an ideal type of bureaucracy and stressed its rational form and organization. In his definition, a bureaucracy is any organization, public or private in nature, that contains seven key attributes. First, the organization must have jurisdiction and be made up of positions that contain detailed responsibilities and scope of authority. Second, there is a hierarchy or a system of supervision and subordination for individuals. Third, there needs to be unity of command and an understanding that although officials do not own the resources, they need to perform their functions so that they are still held accountable for their use. Fourth, bureaucratic organizations must operate on the basis of written documents. Fifth, managers and workers are trained and skilled in the job to assure efficiency and productivity. Sixth, there must be consistent application of rules. Finally, personnel are hired, and work assignments are based on competence and experience. Weber's theory presented an ideal type of bureaucracy for creating the best-running organizations; it was a normative, but not necessarily a descriptive, theory. Weber's bureaucracy emphasized control, efficiency, and productivity.

Hierarchical and bureaucratic structures are now the pervasive form found in business, religious, military, and civilian government organizations today.

The attributes presented in early bureaucracy theory are considered generic organizations—applicable to any form, whether public or private. Although the command and control idea of bureaucracy still exists, more decision-making authority may be more dispersed as policy is achieved in decentralized manners. Ultimately, the structure and functioning of a bureaucracy is related to the organizational and political environment of government.

The Development of American Bureaucracy and Scholarship

Bureaucracy existed long before the founding of America. For example, bureaucracy was present in ancient Egypt and Sumeria, and Confucius was a Chinese bureaucrat. American bureaucracy preceded the founding; for example, Benjamin Franklin developed the post office. Many of the founders were involved in both the politics and the administration of the colonies. The history of bureaucratic theory in America has changed over time. Different principles and themes have emerged and dominated different eras. During the founding years of American government, the bureaucratic theory was government by gentlemen, or aristocratic elites. There was genuine interaction between the executive and legislative branches, and the size of government was small. For example, George Washington had more servants working at Mt. Vernon during his first presidential term than America had bureaucrats.

The terms *administration* and *bureaucracy* are absent from the Articles of Confederation and the Constitution, but early congressional acts formed cabinet departments, appointed officials, and considered administration, or as John Rohr (1986) has argued, administration was an early consideration of the founders. As America grew and changed, the bureaucratic theory of federal operations transformed from an elite corps to a populist notion that anyone should be able to work in the government. Under Andrew Jackson's administration, the spoils system was a theory and practice of patronage appointments as rewards for politics. After the Civil War, the transition from the spoils system to a professional, educated merit system of civil service became the hallmark of the modern administrative state.

There were a number of reform efforts at the end of the 19th century and the beginning of the 20th century that contributed to the development of a professional bureaucracy. The Progressive Era municipal reforms emphasized changing the government structure and operations that would lead to good government; they emphasized improved practices, with professionally appointed administrators, separate from political influence. Scholars were advancing the profession of public administration.

In particular, Frank Goodnow's (1900) textbook argued that the government process is complex and that politics is the expression of the will of the state, whereas administration is the execution of that will. Later, this distinction was called the politics–administration dichotomy, where politicians would make policy, and administrators, as neutral experts, would carry out the policy in an efficient, effective, and economical way. Many of the early textbooks on federal, state, and municipal government emphasized the bureaucratic structures and techniques to accomplish goals. In addition to municipal reforms, another hallmark of the early 20th century was the creation of the merit system for public employees.

The professionalization of bureaucracy was seen as promoting businesslike practices. Woodrow Wilson (1887) is now recognized for his essay in which he stated, "It was harder to run a constitution than to frame one" (p. 200). He argued that a businesslike foundation was needed for government. Wilson admired the British system and was troubled by public influence on elections and government policy; he argued for an expert bureaucracy and a strong executive. Wilson's early writings on bureaucracy became widely known when reprinted in 1941.

During the era of the New Deal, the increased regulation established the national administrative apparatus that we know today. With its expansion in size and scope, there were a number of studies of the federal government to reform the bureaucracy. The Brownlow Commission in 1937 and the Brookings Report in 1938 presented competing views of administration. The Brownlow report argued for an all-powerful executive; the Brookings report argued for more legislative control. Through the passage of the Congressional Reauthorization Act of 1945 and the Administrative Procedure Act of 1946, Congress exerted more control over the bureaucracy. One distinctive theory of this era, Luther Gulick and Lyndall Urwick's (1937) evaluation of the responsibilities of the executive, could be summarized in the acronym of POSDCORB (planning, organizing, staffing, directing, coordinating, reporting, and budgeting). Gulick and Urwick asserted these roles could be delegated by the executive and achieved by top-level bureaucratic leaders.

In the late 1940s, three theorists widened the discussion of bureaucracy from neutral experts to how it affected policy. The New Deal brought the enormous growth of government that continued during World War II. Paul Appleby (1945) noted big democracy combined both policy and administration to achieve public goals; however, contentious questions began to arise regarding the appropriate role of each as agencies—staffed with unelected individuals—began fully gaining power. But the key figure of this era who accepted that administration is politics is Dwight Waldo (1948), whose *Administrative State* makes the case that the purpose of government is to provide for the good life. Bureaucracy can prevent citizens from experiencing life in Hobbes's nasty, brutish, and short state of nature

where the fear of death drives all human motivation. Waldo asserted contemporary citizens do not need to be concerned with the Lockean state of nature where fiduciary trust reigns supreme. In developing a political theory, Waldo shows exactly how politics and administration are interconnected, and this makes him the most influential bureaucracy scholar of the second half of the 20th century.

During the late 1940s through the 1960s, there were a number of alternate theories presented about accomplishing government activities and about the effect of hierarchy and bureaucratic organizations on individual achievement. Herbert Simon (1947) asserted that administrative behavior was productive but it required a type of administrative man loyal to the organization and that his administrative behavior should act in a highly rational manner. In other academic fields, such as psychology and business, theories were expanding views of bureaucratic organizational forms. Abraham Maslow's (1954) hierarchy of needs through which individuals navigate was applied by Douglas McGregor (1960) to two types of organizations and how they treated individuals. According to McGregor's Theory X, management views workers as lazy individuals that need to be micromanaged, and McGregor advocated developing a more flexible Theory Y where hard-working, dedicated individuals were respected by management. Other scholars considered aspects of decision making and group dynamics on governance. Irving Janis (1982) was particularly influential in his description of individuals, finding they will seek unanimity in groups, even if such groupthink has the potential to result in catastrophic government actions.

A final bureaucratic theory of decision making was built on the incremental nature of government. Charles Lindblom (1968) advanced the argument that change in bureaucracy was by increments, not by rational comprehensive studies. Lindblom's suggested bureaucracy worked through evolution, not revolution. Rather than taking big bites, administrators ultimately opt to nibble as they muddle through the decision-making process. Policy change occurs slowly over time. Aaron Wildavsky (1964) applied instrumentalism to budgeting and showed how the best predictor of an agency's next budget was its current budget. These studies help show the importance of policy when examining budgets, bureaucrats, and Congress. Wildavsky also emphasized the obligation of bureaucrats to be proactive and bring up ideas; they must "speak truth to power."

In the last two decades of the 20th century, two competing theories were advanced. First, many were arguing that bureaucracy was undemocratic and it needed to develop better approaches for citizen participation. Second, the new public management theory was advanced in hopes of achieving bureaucratic reform. This theory asserted that good managers were being hampered by excessive rules and red tape. The alternative proposed was for a more innovative, creative, and liberated management

to focus on achieving goals rather than being concerned with structure, process, and constraints of political accountability. As the 21st century began, the bureaucracy was criticized for being uncontrollable and overcontrolled.

The Organization of the Modern Bureaucracy

All levels of government have bureaucratic forms that use agency and departmental structures. This discussion, like many others on American bureaucracy, uses the example of the federal government. Federal bureaucracy is established through the Constitution by Congress enacting legislation that empowers administrative action, while imposing responsibilities and limits. The most visible arm of the federal bureaucracy is composed of cabinet departments, which report directly to the president. Examples include the Department of Education, the Department of Homeland Security, the Department of Energy, and the Department of Justice. These cabinet departments are made up of a variety of agencies that often may have multiple—even conflicting—responsibilities. Consider the Department of the Interior, which contains agencies that deal with both national parks and land management: Parks are to be protected for all citizens to enjoy while land management leases federal property for mining, logging, or ranching purposes. Besides the cabinet departments, a number of other federal agencies independently report to the president—such as the Social Security Administration, the Environmental Protection Agency, and the Office of Management and Budget.

The federal government contains numerous independent regulatory boards and commissions. These are formed through legislation, but these groups do not regularly report to the president. Examples include the Federal Elections Commission, the Securities and Exchange Commission, and the Federal Communications Commission. They have the authority to perform tasks such as setting rates and mandating licensing and registration requirements. The justification for this independence is that boards need to be separated from political control and day-to-day decisions to be effective. These boards are governed by commissioners appointed by the president, with the consent of Congress. The final type of federal bureaucratic organization is the government corporation, which is chartered by Congress. As corporations, they are expected to operate as business enterprises and raise enough revenue to cover their own expenses—though often they do not. Prominent examples include the Tennessee Valley Authority and the United States Postal Service.

Moving from the federal level to the state and local levels, many similarities, along with a few important differences, deserve mention. Although it is not practical in this chapter to discuss the intricacies of every state and local government, it is possible to point to some common elements and some of the most noticeable differences. Many states have cabinet-level departments that are responsible

for actions within each state. Unlike at the federal level, some of the departments are constitutional bodies that have separately elected officials and thus do not report to the governor. For example, positions such as secretary of state, state treasurer, and state auditor are elected in many states rather than being appointed by the chief executive. It ultimately depends on state constitutions and statutes to determine the structure and functions of bureaucracy. The tasks performed by state bureaucrats in many ways mirror the goals of their federal counterparts: They handle issues such as public health and safety, maintaining justice, running elections, and managing finance and budgets.

At the local level, bureaucracy is defined by flatter hierarchies that are affected by the type of government, the population and area to be governed, and the requirements of state constitutions and statutes. Counties, cities, towns, and special districts are all local governments, and the type of bureaucracy will be much different for the City of Chicago than a small rural town in Idaho. Local-level bureaucrats focus on issues such as fire protection, police protection, zoning, and handling issues related to public health. Often, specialized services are provided by special districts devoted to topics such as water treatment or education rather than government corporations or independent agencies. The scholarly literature about the local bureaucrat has considered particular aspects of governing such as mandates from other governments or citizen involvement.

The Critique

Michel Crozier (1964) said bureaucratic organizations are those that demonstrate the slowness, the ponderousness, the routine, the complication of procedures, and the maladapted responses of the agency to its general purpose. Bureaucracy, Crozier noted, is not well liked by the American public. Politicians run entire campaigns against the bureaucracy to win votes. Presidents belittle the role of bureaucrats, such as Ronald Reagan saying that “government is the problem, not part of the solution.” Popular entertainment, the media, and textbooks have all reflected negative views of civil servants over the last four decades. These often represent government bureaucracy as all-powerful and out of control.

Bureaucracy is often associated with negative phrases like red tape, inefficient, duplicative, big, unresponsive, inhuman, inflexible, machinelike, costly, out of touch, rule bound, lazy, power hungry, uncontrollable, monolithic, secretive, wasteful, and unrepresentative. There is a long tradition of highlighting the negative aspects of bureaucracy. Ralph Hummel (1977) points out that the bureaucratic form leads to problems and that bureaucrats are often misunderstood. In considering how bureaucracy shapes the public, Hummel presents pathologies of the bureaucratic structure, including that bureaucracy deals with cases, not people, and dehumanizes its clientele. Bureaucrats worry about control and efficiency rather than justice, freedom,

and salvation. Instead of speaking our language, headless, soulless bureaucrats create their own secret languages while working for control institutions, rather than service institutions.

What these critiques overlook is the essential role that bureaucracy plays in the proper functioning of government. As the executive branch, bureaucracy serves as the executor of the decisions of legislative bodies. It implements policy and takes it from paper to action. The American tradition neglects such positive factors and instead encourages public and political skepticism about what, how, and why bureaucracy does things. Bureaucrats are accused of poor performance and self-interest. Citizens and academics criticize their personalities, communication skills, and timeliness in response. Normal bureaucratic accomplishments and functions are boring for media reports; consequently, the stories presented on national and local news accentuate the negative to highlight bureaucratic ineptitude and waste. The general distrust about the nature of government ends up as bashing the bureaucracy.

There are many more scholars that have studied bureaucratic problems than the few individuals that have a more positive outlook. Charles Goodsell's (1983) polemic focused on the personal experience of an individual's encounter with a bureaucratic agency and found that bureaucracy performs better than it is given credit for; individuals were pleased with their experience with public servants. Richard Stillman (1996) carefully explains away the 10 biggest myths about bureaucracy. He makes the case that bureaucracy is not the problem in American government, bureaucracy is not overwhelmingly large or monolithic, bureaucrats are not all alike, bureaucrats do not remain in their positions forever, bureaucracy is not entirely concentrated at the federal level, bureaucracy does not operate in secret, and bureaucracy is not wasteful, all-powerful, or nothing more than red tape.

Applications and Empirical Evidence

Studies of the bureaucratic institutions of government use many different research approaches, including empirical positivist studies, qualitative studies, and in-depth field studies. The methods used often include a combination of approaches or the application of research triangulation. Often with studies of the bureaucracy, the evidence gathered relates to the type of inquiry. Studies of management and performance, use of resources and budgets, and program evaluation involve positivist and empirical approaches. Studies of policy adoption, implementation, administrative politics, and interactions between branches may use qualitative or mixed techniques. Surveys of public opinion about government activities and bureaucracy are also common.

The application of theory through systematic, data-driven studies was first adopted for management and practical

purposes. Frederick Taylor's (1911) scientific management theory encouraged systematic study of how work is conducted to improve operations. It is fair to say that scientific management, with its emphasis on efficiency and management, was a worldwide movement that had enormous influence on business practices and government policy. The scientific management approach relies on gathering information, using time and motion studies, evaluating tasks, and evaluating quantitative data. Managers use these studies to determine the best techniques and the appropriate division of work. These studies were often agency based; sometimes they dealt with policy considerations.

Many early works discussing bureaucracy were normative based. The use of a hierarchical structure, with clearly defined roles and responsibilities, was seen not just as a descriptive theory, but also as a practical application to exert control over workers in the bureaucracy. The approach to bureaucratic studies during the first half of the twentieth century was applied research used to improve practices. The resulting studies were descriptive studies and theories of organization. By the 1940s, the terms shifted from bureaucratic or organizational studies to administrative studies. These studies tended to be in substantive policy areas, but some also occurred within a theory and knowledge-building framework, toward a more academic and systematic type of policy studies.

Herbert Simon (1947) presented a pointed critique of these bureaucracy studies that, in his view, provided contradictory proverbs in their approach rather than a valid analytical administrative theory. Simon presents an alternative to develop a systematic scientific administrative theory. Simon's rational comprehensive approach to the study of administration is similar to positivist research. The first requirement was consistent and discrete definitions of concepts that enable studies of administrative situations to be comparable. Second, scholars needed to use rigorous social scientific methods. Simon proposed highly rigorous studies to test and develop theory including the following: falsifiable hypotheses, discrete and measurable criteria, common administrative situations, and replicable studies. Simon argued for experimental designs where possible. Simon's ideal was to develop a theory of administration built on exhaustive and rigorous study, but he recognized that there were limits on the time, money, and ability to conduct comprehensive studies. As a result, many studies consider satisfactory and sufficient, even if limited, information. In developing and conducting satisfactory and sufficient information, the researchers "satisfice."

In 1951, the policy sciences approach from Daniel Lerner and Harold Lasswell (1951) shifted the study of bureaucracy from organization and management to considering government policy. They asserted there was a scientific discipline in the study of public policy. Policy sciences was proposed as a better way to study and understand

public policy, and the bureaucracy was seen as the setting or a part of the inquiry rather than bureaucracy being the primary focus. The policy sciences and administrative theory share an approach that uses the scientific method, hypothesis testing, and choice of problems that are scientifically tractable, and they also preferred applying statistical analysis and mathematical modeling.

In recent years, the move from understanding a particular substantive area of bureaucracy has been supplemented with a call for developing a better theory for understanding bureaucracy's role in the policy process. Again, one aspect of this policy inquiry is the study of how bureaucracy affects policy. Paul Sabatier (1999) argues that multiple method frameworks that consider quantitative data, longer time frames, and criteria across levels of government or comparative between countries must be developed to build better theory.

There are a number of studies that test theories of bureaucracy. The use of a hierarchical structure, with clearly defined roles and responsibilities, was seen not just as a descriptive theory, but also as a practical application to control the structure, functions, and activities of the bureaucracy. Other applications of theory are related to practical considerations of describing and evaluating bureaucratic practices. A common inquiry using empirical evidence is through survey and field study research. Numerous internal surveys of bureaucrats have been conducted. These consider what bureaucrats do; their opinions about their tasks, responsibilities, and processes; and their view of individual or agency accomplishments. Some surveys focus on particular levels of the organizations; others may be cross-sectional in comparing agency types across states or nations. It is more common to find single-shot surveys than the use of longitudinal studies. Other studies may inquire as to the power, process, implementation, or relationships within agencies as these relate to policy areas. Common methods for these research questions are in-depth field studies and observations.

Some studies will look at how a particular program is operating. This is particularly true of the implementation studies. In the 1970s, two studies, by Jeffrey Pressman and Aaron Wildavsky (1973) and Eugene Bardach (1977), considered how bureaucracy takes legislation and implements policy, how policy changes, and what problems occur. These implementation studies tend to be case studies and often consider both policy purposes and bureaucratic processes. Another area of inquiry about bureaucratic achievement is in the area of evaluation studies. Again, the focus is often on one program, but the design often includes quantitative data analysis to address questions that compare policy purposes to policy goals. Evaluation studies also consider the effects of policy on the targeted population to develop better policy tools for bureaucrats to affect behavior.

Inquiries about the structure and functions of various agencies, including how they define their policy responsibilities,

may be conducted using documentary analysis. Numerous approaches and sources for documentary analysis may begin with the policy delegation through statutes, considering internal memoranda and documents, or through analysis of rule-making or other standard operating procedures. At the federal level, the government manual provides brief summaries of the department's policy role and responsibilities that demonstrate hierarchical structure and designated functions. There is equivalent information available now on government websites for many state agencies.

The use of systematic empirical studies to understand and improve the performance of bureaucracy has an established history. The initial purposes of these studies were for management improvement rather than academic inquiry. These systematic evidence-based studies emphasized the use of data about bureaucracy as a way to understand and control administrative practices. Many of these studies and methods were related to government reform efforts, and the approaches were applied to governments at all levels. From Taylor's (1911) scientific management to the development of program budgeting, the linking of systematic gathering and analysis of performance has had practical use and has expanded knowledge. The most recent systematic performance legislation, the Government Performance and Reporting Act of 1993, requires agencies to gather performance information and report it to both the executive and to Congress. It serves both management and democratic informational purposes.

Empirical studies are also conducted on primary and secondary source documents. Bureaucracy produces many reports as part of normal operations; these are primary source documents, some of which are published, while many others can be gathered through freedom of information applications. There are often evaluations of agencies and programs conducted by an internal auditing agency, include the Government Accountability Office and the agency's inspector general. Other sources of information about bureaucracy include lawsuits and judicial decrees. These numerous documentary sources provide data for evaluation of bureaucracy and policy and often the research design for secondary sources uses qualitative rather than quantitative methods of analysis.

An essential concern about bureaucracy is the influence bureaucracy has on policy making and how unelected bureaucrats use their power. The relationship among bureaucrat, legislature, and interest group has been studied through case studies and field research. Because of the type of interactions, the use of observations, interviews, and documentary analysis is seen as a worthwhile approach to understanding these relationships. The power and influence of bureaucracy on the general public or affected groups is also a topic of note. It is common to study public opinion about bureaucrats, especially on citizen interactions with bureaucracy.

Policy Implications of Bureaucracy

The role of the large unelected institution, the bureaucracy, and its effects on policy has been a persistent concern. Common questions investigate how bureaucrats affect agenda setting, policy selection, decision making, and implementation. In the formal view of government, the three constitutional roles of the judiciary, the executive, and the legislature are defined, but the influence of the bureaucratic institutions of big government on policy is not part of the founding framework. Some have called bureaucracy the fourth branch. Policy concerns relating to the form and structure of bureaucracy often consider how bureaucracy is accountable to other government institutions and the public interest. The effect of bureaucracy on the public interest involves both democratic and political accountability. With the enormous size of the administrative state, there are concerns at the federal level about the ability of the president to control the bureaucracy. There are also questions about whether the bureaucracy is faithful to legislative intent in implementing policy decisions.

The concerns about accountability raise questions about delegation: whether the legislature provides narrow mandates or broad discretionary power to agencies to carry out policy. Theodore Lowi (1969) recommends that the legislature provide narrow and strict mandates as a way to limit the authority of bureaucrats; this is called a top-down approach. Other policy scholars discuss different approaches to delegation based on the type of policy consideration. Anne Schneider and Helen Ingraham (1997) assert that broad administrative discretion is a beneficial, pragmatic approach to achieve policy because it provides necessary flexibility, respects agency expertise, and includes the public. This also reflects a bottom-up approach where granting broad policy discretion includes the public in policy development and implementation.

The problem of delegation and agency accountability is sometimes considered in relation to the principal-agent theory adapted from the fields of economics and law. Where a principal, here a government, relies on agents or employees to carry out responsibilities, there are problems of information asymmetry, agent self-interest, monitoring activity, and moral hazard. The agent has more access to information than the principal, and this asymmetry affects the principal's ability to decide and achieve policy. The problem of self-interest asserts the agent seeks his own interest, which leads to inefficiencies, fraud, or waste. The moral-hazard problem addresses making poor economic decisions because someone else bears the costs. The need to monitor by the principal has to balance costs and control. These concerns with principal and agent relate to how policy is achieved.

David Weimer and Aidan Vining (2005) summarize how policy is evaluated in a political-economy framework. Economic theory asserts there are limits to government intervention (referred to as government failures) in a free

market. This government-failure framework posits reasons why policy cannot be achieved because of the limits of government in influencing political or economic behavior to achieve public purposes. The four categories of problems are direct democracy, representative government, bureaucratic supply, and decentralization. The first two categories relate to government structure in a constitutional democracy; the final two categories specifically relate to bureaucracy and policy. The bureaucratic supply problems include the principal-agent problem, problems of measurement or the need to impute value to output, the problem of limited competition in valuing government services provision, and the inflexibility of civil services provisions. In short, bureaucratic failure is equated to a type of market failure. Decentralization problems also relate to bureaucracy in the areas of inconsistent implementation and fiscal externalities.

An established area of policy inquiry focuses on bureaucratic culture and autonomy. The delegation of authority and the inevitable specialization encourages agencies to act as a closed group of experts. In addition, the dominant hierarchical structure and development of agency cultures may work to perpetuate agency goals as opposed to policy goals of elected officials. The different types of roles and responsibilities given to agencies in carrying out policy, from enforcement to rule making, have produced a powerful administrative state that needs political controls. That is, with the expansion of the administrative state in the first half of the 20th century, the inconsistent and confusing administrative procedures to develop and enforce statutes and regulations became an issue for administrators and lawyers. The adoption of the Administrative Procedure Act led to uniform standards in rule making, imposed due process requirements in agency hearings or adjudication, and required more transparency through public notices and open meetings. There were numerous overhead legislative actions during the second half of the 20th century to make the bureaucracy subject to and responsive to legislative supervision. David Rosenbloom (2000) asserts this legislative-centered focus made the administrative state more amenable to the constitutional representation of the public interest.

Individuals and groups affected by bureaucratic practices can make policy challenges and accomplish substantial policy change through the use of litigation. Through the use of class-action and public-interest litigation, policies in the area of racial discrimination, institutionalization of the mentally ill, prison practices, and other issues were successfully changed. Not only was the bureaucracy subject to review and supervision by the executive and legislature, but bureaucratic actions were considered in the courts with far-reaching influence on policy and bureaucratic practices.

The ability of the agencies to influence policies, agenda setting, decision making, and implementation has been an area of attention. How individual agencies interact with

legislative oversight committees and public interest groups has been considered. The particular power relationships in these three groups combine to affect or control policy in a way that is often called an iron triangle or a subgovernment. A somewhat broader concept of the power, expertise, and influence of legislative experts, public interest groups, bureaucrats, and the public is called a policy subsystem or policy network, and this concept is used to consider changes in substantive areas of policy. Iron triangle, or subgovernment, analysis is focused on how elite groups affect policy; the subsystem or network approach incorporates a pluralist framework that sees bureaucracy as one of a number of actors in the policy cycle, particularly on agenda setting or decision making. These analyses about policy actors and their influence include concerns about the independence, control, and responsiveness of the bureaucracy.

One last area of consideration of the interaction of policy and bureaucracy is in implementation. There are concerns about policy slippage between the adoption of policy and the process of implementation. Some suggest this is a normal result from the broad, general, and optimistic policy preferences, and with the institutional limits of time, resources, and attention, the changes from policy aspirations and bureaucratic development of procedures and processes to implement the policy are part of a normal process. Some scholars have suggested better policy design, and others have critiqued the idea of automatic implementation of policy. Others have noted that there are progressions of policy, and often there are contradictions and paradoxes throughout the policy process that become evident during implementation and evaluation.

Future Directions

There are important questions for political scientists to investigate regarding bureaucracy and how this institution accomplishes government policy and goals. This section considers four broad research areas: institutional organization and reform; government performance and productivity; political responsiveness and bureaucratic power; and democratic values, public service, theory building, and empirical studies.

The theoretical foundations of bureaucratic study emphasize the structure, functions, and control of the administrative state. Questions regarding this important institution and how it can be reformed are enduring. How do we develop better bureaucracies? What can be done to improve or supplant the rigid hierarchical structures? In service delivery, can networks, collaborative efforts, or coproduction produce good results? The use of contracting out and collaborative efforts has been advocated as a flexible and effective way, an alternative to bureaucratic production, to provide government services. However, empirical studies that demonstrate the achievement of this

theory have been lacking. Donald Kettl and Brinton Milward (1996) have criticized this approach as unproven, and they call it the hollow state. The collaborative efforts between governments and nonprofits in achieving public policy is another area for inquiry: Does it work?

The dominant administrative values of neutral and efficient delivery of services have their roots in the Progressive Era. However, questions endure regarding government performance and productivity. Are these the appropriate goals to pursue? Do they supplant or support political goals and the public interest? Does the fiscal and budgetary pressure on bureaucrats to be productive raise different types of questions about political responsiveness? Does the emphasis on productivity and achieving goals lead to behaviors that encourage creativity and risks?

There are a number of enduring questions about political responsiveness and bureaucratic power. Are bureaucrats responsive to the policy goals delegated to them by the legislature? Are they creatures of the executive? Do the courts serve as controls on bureaucratic excess, or do they reinforce bureaucratic power? The concerns about a vast institution that is not elected and has only indirect political controls are raised in many instances. Does the power of the bureaucracy respond to constitutional designs of checks and balances? Are there better models of policy implementation and evaluation that will increase political responsiveness? Are bureaucrats more responsive to interest groups or personal interest than to policymakers?

The democratic values of citizen participation and the public interest are an important part of the American polity. According to William Gormley and Stephen Balla (2004), both scholars and the public believe bureaucracy to be undemocratic. How do individuals interact with the bureaucracy, and how does this affect public perceptions of government? The role of bureaucrats in teaching, encouraging, and engaging the public in policy development and service delivery is proposed by Cheryl King and Camilla Stivers (1998). Does this approach promote or impede the general public interest? These scholars raise the question of how to make bureaucracy more democratic. One additional consideration about democracy and the public interest is the composition of the civil service. Many suggest that there should be a representative bureaucracy that can promote and respect the diverse populations and provide better service. This raises questions of how to evaluate this assertion.

Finally, bureaucracy is often seen as a profession where theory building and other studies predominately are concerned with practical results. How can political scientists develop rigorous theories and use sophisticated methods to predict and understand governance? Are there ways to develop longitudinal studies, rather than single-shot or case studies? How can studies of public perceptions about bureaucracy lead to better government? What ways

are there to further advance to the challenge of Simon's administrative theory or Lerner and Lowell's policy sciences?

Conclusion

Bureaucracy is the institution of government that accomplishes government policies and goals that are established by the legislature and the executive. The theories of bureaucracy have changed as this nation has developed, but the dominant form of hierarchical control remains in use. The application of bureaucratic form and structure are part of the theories that seek to control the scope and complexity of the administrative state. The instrumental and practical concerns about structure, control, and efficiency of the bureaucracy often compete with important democratic values of political responsiveness, policy influence, and public interest. There are important questions that 21st-century scholars need to continue to study to understand how bureaucracy does or does not achieve political goals.

References and Further Readings

- Appleby, P. H. (1945). *Big democracy*. New York: Knopf.
- Bardach, E. (1977). *The implementation game: What happens after a bill becomes a law*. Cambridge: MIT Press.
- Crozier, M. (1964). *The bureaucratic phenomenon*. Chicago: University of Chicago Press.
- Goodnow, F. (1900). *Politics and administration: A study in government*. New York: Macmillan.
- Goodsell, C. T. (1983). *The case for bureaucracy: A public administration polemic*. Chatham, NJ: Chatham House.
- Gormley, W. T., & Balla, S. J. (2004). *Bureaucracy and democracy: Accountability and performance*. Washington, DC: CQ Press.
- Gulick, L., & Urwick, L. (Eds.). (1937). *Papers on the science of administration*. New York: Institute of Public Administration.
- Herring, E. P. (1936). *Public administration and the public interest*. New York: McGraw Hill.
- Hummel, R. (1977). *The bureaucratic experience*. New York: St. Martin's.
- Janis, I. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*. Boston: Houghton Mifflin.
- Kettl, D. F., & Milward, H. B. (Eds.). (1996). *The state of public management*. Baltimore: Johns Hopkins University Press.
- King, C. S., & Stivers, C. (1998). *Government is us: Public administration in an anti government era*. Thousand Oaks, CA: Sage.
- Kingsley, J. D. (1944). *Representative bureaucracy: An interpretation of the British civil service*. Yellow Springs, OH: Antioch Press.
- Krislov, S. (1974). *Representative bureaucracy*. Englewood Cliffs, NJ: Prentice Hall.
- Lerner, D., & Lasswell, H. (Eds.). (1951). *The policy sciences: Recent developments in scope and method*. Stanford, CA: Stanford University Press.
- Lindblom, C. E. (1968). *The policy making process*. Englewood Cliffs, NJ: Prentice Hall.

- Lowi, T. (1969). *The end of liberalism: Ideology, policy, and the crisis of public authority*. New York: W. W. Norton.
- Maslow, A. (1954). *Motivation and personality*. New York: Harper & Row.
- McGregor, D. (1960). *The human side of enterprise*. New York: McGraw Hill.
- Meier, K. J., & O'Toole, L. J. (2006). *Bureaucracy in a democratic state: A governance perspective*. Baltimore: Johns Hopkins University Press.
- Milward, H. B., & Provan, K. G. (1993). *The hollow state: Private provision of public services*. Washington, DC: Brookings Institute.
- Pressman, J., & Wildavsky, A. (1973). *Implementation: How great expectations in Washington are dashed in Oakland*. Berkeley: University of California Press.
- Rohr, J. (1986). *To run a constitution: The legitimacy of the administrative state*. Lawrence: University Press of Kansas.
- Rosenbloom, D. (2000). *Building a legislative centered public administration: Congress and the administrative state, 1946-1999*. Tuscaloosa: University of Alabama Press.
- Sabatier, P. (1999). *Theories of the policy process*. Boulder, CO: Westview Press.
- Schneider, A. L., & Ingraham, H. (1997). *Policy design for democracy*. Lawrence: University Press of Kansas.
- Selden, S. C. (1997). *The promise of representative bureaucracy: Diversity and responsiveness in a government agency*. Armonk, NY: M. E. Sharpe.
- Simon, H. A. (1947). *Administrative behavior: A study of decision making processes in administrative organization*. New York: Macmillan.
- Stillman, R. E., III. (1996). *The American bureaucracy: The core of modern government*. Chicago: Nelson Hall.
- Taylor, F. W. (1911). *The principles of scientific management*. New York: Harper & Brothers.
- Terry, L. D. (1995). *Leadership of public bureaucracies: The administrator as conservator*. Thousand Oaks, CA: Sage.
- Waldo, D. (1948). *The administrative state: A study of the political theory of American public administration*. New York: Ronald.
- Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (G. Roth & C. Wittich, Eds.). Berkeley: University of California.
- Weimer, D. L., & Vining, A. R. (2005). *Policy analysis: Concepts and practices*. Saddle River, NJ: Pearson.
- Wildavsky, A. (1964). *The politics of the budgetary process*. Boston: Little, Brown.
- Wilson, W. (1887). The study of administration. *Political Science Quarterly*, 2, 197-222.

87

INTEREST GROUPS AND PLURALISM

STEPHEN SUSSMAN

Barry University

An interest group can be defined as an organized group sharing common objectives that actively attempts to influence government (Janda, Berry, & Goldman, 1997). Interest groups are private organizations that try to affect public policy and try to influence the behavior of political decision makers. As a result, interest groups are often called pressure groups because of their effort to exert pressure in an effort to promote their agenda. The term *interest group* covers just about any group of people attempting to influence government. David Truman (1951) defines an interest group as “a shared-attitude group that makes certain claims on other groups in society” (p. 37) by acting through the institutions of government. Some interest groups are temporary; others are permanent. Some focus on influencing a particular policy, others on broad changes. Some work through the executive or administrative agencies, others through the judicial or legislative sectors, still others through public opinion.

Membership in interest groups is linked to socioeconomic status and is greatest among professional, college-educated, and high-income persons. Those who are the most disadvantaged economically cannot afford to join interest groups, and many may not have the time or expertise to find out what group might represent them. Even a small contribution to an interest group may be seen as a luxury they cannot afford. Nor do they have the expertise to be interest-group entrepreneurs by creating and marketing their own special interest. As reported in 1993, although

25% of social security recipients were members of an interest group that supported the needs of social security members, less than 1% of the food stamp recipients and only 2% of recipients of Aid to Families with Dependent Children were members of groups that supported their interests (Verba, Schlozman, Brady, & Nie, 1993).

Where do interest groups fit in the policy cycle? Government is defined as the official policymakers. Official policymakers have the legal authority to make policy. The legislative, executive, and judicial branches of government make up the body of official policymakers. Interest groups are considered part of the nongovernmental participants. Nongovernmental participants have no legal authority to make policy; they can only influence policy. Interest groups, political parties, and the media are examples of nongovernmental participants that can affect policy but not make policy. Political parties and the media can be included as interest groups. Political parties, like interest groups, try to influence public policy. However, interest groups can support candidates for public office, while political parties sponsor candidates for public office. Interest groups may influence the nomination of candidates. However, political parties must win elections. Candidates do not run for office under the banner of the interest group. However, they do run under the banner of the political party.

There are many types of interest groups that can affect policy as nongovernment players. They include corporations,

trade associations, professional associations, labor unions, citizen groups (interest groups open to anyone), think tanks (Heritage Foundation, Cato Institute, Rand Corporation), universities and colleges, churches, charities, media outlets, and political parties. These groups are part of the agenda-setting, policy formulation, and evaluation process. Their power and authority come from their membership and financial status. The goal of the interest group is good public policy, but individual members may have self-interests.

The First Amendment to the Constitution gives every citizen the right to assemble and petition the government for the resolution of grievances. Alexis de Tocqueville (1834–1840) wrote, “In no country of the world has the principle of association been more successfully used, or more unsparingly applied to a greater multitude of objects, than in America” (p. 197). He was amazed at the degree to which Americans formed groups to solve social and political problems and suggested that the ease with which Americans form organizations is a reflection of a strong democratic culture.

James Madison (1787) viewed interest groups, which he called factions, as a necessary evil. He believed that interest groups not only conflict with each other but also conflict with the common good. However, Madison thought that abolishing interest groups would destroy liberty—a remedy “worse than the disease.” In “Federalist No. 10,” published in 1787, he explained that liberty promotes factions, and in a large republic such as the United States, there would be many different factions held together by regional or local special interests so that no single one of them would dominate national politics.

Use of the term *special interest* implies that it is not in the public interest. Many scholars dislike the term, since it can carry a negative implication. When politicians want to garner support from constituents, they may refer to the opposition as being friendly to special interests. This suggests that the special interest does not have the public interest at heart. Public interest groups are citizen groups that have no economic self-interest in the policies they pursue (Berry, 1996). Groups such as Common Cause and the Consumer Federation of America see themselves as public interest groups, interested in the public welfare and not the interests of business organizations, trade associations, unions, or other special interests. They want government to take a largely regulatory role. As a result, public interest groups are popular with environmental and civil rights groups, both of which serve special interests.

Government growth and regulation over the last few decades account for a good portion of the proliferation of special interests. The more areas in which the federal and state governments have become involved, the more special interests have developed to attempt to influence policy. Veterans’ benefits create veterans’ groups; professional license requirements by state governments have given rise to groups such as the American Medical Association, the

American Bar Association, and other groups that lobby government for their own special interests. Interest groups have also increased over the last few decades as government becomes more active in civil rights, social welfare, and consumer rights.

There are many interest groups looking to influence government to benefit their special interest. They include business and trade organizations, professional associations, organized labor, women’s organizations, religious groups, single-issue groups, and environmental organizations, among others.

Business and trade organizations include groups such as the U.S. Chamber of Commerce and the National Association of Manufacturers. Professional associations include groups such as the American Bar Association and the American Medical Association. The Teamsters and the American Federation of State, County, and Municipal Employees (AFSCME) are examples of organized labor. Women’s organizations date back to the pre-Civil War period when women were active in the antislavery campaigns. In the early 20th century, women’s organizations were active in the suffrage movement. Today, the largest feminist organization is the National Organization for Women. Religious groups have a long history of involvement in politics. The Southern Christian Leadership Conference took an active role lobbying for the Civil Rights Act of 1964. Other politically active groups include the Christian Coalition and the Anti-Defamation League of B’nai B’rith. Single-issue groups, such as the National Rifle Association or Mothers Against Drunk Driving (MADD), attract support from people with a strong commitment to a single cause. Groups such as the National Wildlife Federation, the Sierra Club, and the National Audubon Society lobby to protect the environment.

James Madison’s solution to the problems posed by interest groups was to create a wide-open system in which many groups would be able to participate so that groups with opposing interests would counterbalance one another. He succeeded in his efforts. Today, some scholars argue that the increase in interest group representation has resulted in less power for interest groups and better democracy.

Gabriel Almond (1958) identified four types of interest groups: anomic, nonassociational, institutional, and associational. Anomic groups are generally spontaneous groups with a collective response to a particular frustration. Anomic groups tend to generally have poor political communication skills. Nonassociational groups are rarely well organized, and their activity is dependent on the specific issue. They differ from anomic groups in that members are usually similar to one another and have a common identity. Institutional groups are mostly formal and have some other political or social function in addition to the particular interest. Associational groups are formed specifically to represent an issue of a particular group. The institutional category includes legislatures, political executives,

bureaucracies, armies, and churches. These groups articulate their own interests or represent the interests of other groups in the society. The nonassociational interest groups include lineage, ethnic, regional, religious, status, and class groups. The anomic category is concerned with spontaneous breakthrough into the political system, such as riots and demonstrations. The associational interest groups include specialized groups such as trade unions, business organizations, professional organizations, civic organizations, and educational organizations. They form to represent the interest of a group.

Public interests are interests that are connected to the general welfare of the community, while private interests are associated with benefits for part of the community. Political parties offer people the best chance to participate in the decision-making process (Schattschneider, 1960).

Private interest groups include producers, professional groups, and unions. Producers represent goods or services, such as business and agriculture. Professional groups represent the interests of professionals, such as doctors, lawyers, and dentists; they are very influential in the policy-making process because of their importance in local communities and because of their ability to make substantial campaign contributions. The unions' primary role has been to protect the jobs of their members and work for maximum wage and benefit levels, such as pensions; political power of private-employee unions has eroded, possibly due to declining membership. In general, labor union membership has always been low in the United States compared with other Western nations, and it has been steadily declining in recent years; most observers believe that the political power of private labor unions has eroded in dramatic ways over the past several decades.

Public interest groups try to get government to do things that will benefit the general public rather than the direct material interests of their own members. The number and influence of public interest groups has grown markedly since the late 1960s; many were initiated by social movements, including some that were created by social movements for consumers, the environment, and women.

The Bureau of Labor Statistics shows that private-sector union membership has steadily declined since 1973. At the same time, public-employee union membership has steadily increased. The *Wall Street Journal* reports that in 2009 the number of unionized workers who worked for the government surpassed those in the private sector for the first time. The Bureau of Labor Statistics reported that 51.4% of the 15.4 million union members in the United States were employed by government in 2009 (Stern, 2010). In 1962, President Kennedy signed executive order 10988, allowing the federal work force to unionize. This led to the rise of the unionized public work force, and it effectively changed the political environment. Unions such as AFSCME, the Service Employees International Union, and the National Education Association had incredible growth (Henninger, 2010). Since public-sector

salaries come from taxpayers, these public-sector unions are representing the interests of their members when they advocate for higher taxes.

Mancur Olson (1965) focused on the logical basis of interest group membership and participation. He theorized that only a benefit reserved strictly for group members will motivate one to join and contribute to the group. This means that individuals will act collectively to provide private goods but not to provide public goods. In 1982, he expanded the scope and suggested that that small coalitions tend to form over time in countries. Groups and unions will have incentives to form political lobbies and influence policies in their favor. These policies will tend to be protectionist and antitechnology and will therefore hurt economic growth. It can be argued that protectionist and antitechnology policies help maintain the status quo, thereby protecting jobs. This generally does not work and can lead to economic decline (Olson, 1982).

Pluralism Model

The political theory of pluralism holds that political power in society does not lie with the electorate, or with a small concentrated elite, but is distributed among wide numbers of groups. These groups may be trade unions, interest groups, business organizations, or any of a multitude of formal and informal coalitions.

Most political scientists in the United States believe that interest groups actually serve as important instruments to attain democracy and serve the public interest. Pluralists believe the interest group system is democratic because people are free to join or to organize groups that reflect their own interests. Generally, elections do not adequately communicate what the people want in terms of policy; the many groups and organizations to which people belong are better vehicles to convey what the people want to political leaders. Groups are easy to form. Most citizens are able to create them without too much difficulty. As a result, the interest group system serves the U.S. system well because people in the United States are free to join or to organize groups that reflect their interests, and it provides a vehicle for citizens to influence government. Government power is broadly dispersed because of federalism, checks and balances, and the separation of powers; because power is broadly dispersed, governmental institutions are open to the appeals of the broad variety of groups that exist in society.

Pluralist political scientists view interest groups as an additional tool of democratic representation (and not as a problem). The pluralist model is the prominent theory associated with interest groups. In this theory, democracy is viewed as a marketplace with more or less perfect competition. Various perspectives are represented by interest groups that compete for influence over policy issues. The pluralist model assumes equal access to the policy-making arena, fragmentation of the marketplace, a competitive

process for determining policies, and the neutrality of government. These theoretical conditions form a system in which everyone is free to organize for political purposes and in which the policy-making process is not monopolized by powerful political forces.

All Western democracies exhibit some degree of pluralism, but the United States may be the closest to a realization of this model. However, the pluralist model has been criticized on several accounts. The predominant argument has been that in reality, political access and power are unevenly distributed in the democratic system as a result of varying levels in education, economic resources, and political constraints. Early criticism of pluralism led to the development of the elitist theory, which states that financially privileged individuals and groups have more impact on policy making than other groups.

Traditionally, interest groups have been seen as special interests that challenge the public good and make the pursuit of the public interest extremely difficult. By contrast, pluralists believe that interest groups are democratic because people are free to join or to organize groups that reflect their own interests. Pluralists see interest groups naturally forming when people are adversely affected. The pluralist model of democracy is based on the idea of pluralism, which assumes that people form groups along economic, religious, ethnic, or cultural lines and attempt to influence government policymakers. The pluralist model of democracy interprets government by the people to mean government by people operating through competing interest groups.

Pluralism views governmental conflict in terms not of a majority versus a minority but of many minorities in conflict with one another. Pluralism views society in terms of a conflicting struggle among different interests. Pluralist theory includes the belief that new interest groups form when the need arises. Truman (1951) stated that when individuals are confronted with change, they group together in an interest group. He viewed interest groups as a self-correcting market force. No government entity forces interest groups to form. They form to remedy some imbalance in society.

Under pluralism, democracy comes about through the openness of the system to group interests, but not necessarily as a result of mass citizen participation. Dahl (1989) stated that no modern country meets the ideal of democracy, which is as a theoretical utopia. To reach the ideal requires meeting five criteria: (1) Effective participation—citizens must have adequate and equal opportunities to form their preference and place questions on the public agenda and express reasons for one outcome over the other; (2) voting equality at the decisive stage—each citizen must be assured his or her judgments will be counted as equal in weight to the judgments of others; (3) enlightened understanding—citizens must enjoy ample and equal opportunities for discovering and affirming what choice would best serve their interests; (4) control of the agenda—people must have the opportunity to decide what

political matters actually are and what should be brought up for deliberation; and (5) inclusiveness—equality must extend to all citizens within the state. Everyone has a legitimate stake within the political process. Pluralists emphasize that power is not a physical entity that individuals either have or do not have but flows from a variety of different sources. Power comes from controlling various resources. However, because society has many interests, under pluralism, the common good is best served by a process that permits many interests to achieve their policy goals. If many interests are served, then the collective interests of society will have been served (Dahl, 1961).

Elite theory is a theory that tries to describe and explain the power relationship in the political environment. It argues that economic and policy-planning power is held by a small minority. Through positions in corporations, on corporate boards, or through financial support of for-profit and not-for-profit organizations, members of the elite are able to exert significant power over policy decisions of corporations and governments. Although there are some objections to the elite model of pluralism, groups often need financial and political resources to effectively operate. “The flaw in the pluralist heaven is that the heavenly chorus sings with a strong upper-class accent” (Schattschneider, 1960, p. 35). Essentially, money may be the single most important factor in interest group success. Money is especially important for elections, and groups help candidates who favor their cause.

The Purpose of Interest Groups

Interest groups serve at least five purposes in American government. Interest groups represent their constituents in an effort to affect the political agenda. Interest groups make it possible for citizens to participate in the political process. They educate group members, the general public, and politicians. They help bring attention to the specific issues, and they help oversee and evaluate current programs. As program evaluators, they act in an unofficial capacity, making sure their interests are served. Although the federal government has the Government Accountability Office, the Office of Management and Budget, and other official evaluators, interest groups also evaluate programs. Interest groups may help to get an issue on the legislative agenda, they may help an agency write a rule, and they may bring a case to the U.S. Supreme Court. However, interest groups will do unofficial evaluations to determine success of a policy and determine if the group needs to lobby for more change.

Membership in interest groups generally requires dues. Many interest groups have strong member bases and huge financial resources, giving them frequent access to public officials. Business and professional organizations tend to have little problem maintaining a membership. Members of these groups gain a substantial benefit by joining. The American Bar Association, the American Medical Association, and the U.S. Chamber of Commerce are

examples of business and professional organizations with a strong membership. As a result, they employ very powerful lobbyists and have little problem gaining access to government.

Other groups serve an ideological purpose. These interest groups tend to have members who drop in and drop out of the group. Large citizens' groups, such as the American Association of Retired Persons (AARP), MADD, and the NAACP, have huge financial resources because of their extensive membership. Although members may drop out, others are always joining. As a result, they can hire very effective lobbyists. Smaller interest groups, with limited financial resources, may not be able to afford an effective lobbyist.

Lobbyists

In addition to membership, interest groups employ lobbyists. Lobbyists represent the interest group and promote the group's interest to the official policymakers. Lobbyists can devote their efforts to one organization or interest, or they can be part of a law or consulting firm. The term originated at a time when interest groups tried to contact members of Congress in the lobby of the Capitol building.

Many lobbyists are lawyers and former government officials who are familiar with the political environment. The growth of the professional lobbyist has meant a concentration of power over government within an elite group of people, with personal and professional connections that help gain them access to policymakers and legislators.

The two main lobbying strategies are labeled *inside lobbying* and *outside lobbying*. These strategies include communication between public officials and group lobbyists. Inside lobbying is based on developing close contacts with policymakers. Outside lobbying uses public pressure to influence officials (Ornstein & Elder, 1978).

There are two broad categories of lobbying: direct lobbying and grassroots lobbying. Direct lobbying requires direct contact with the policymaker (legislative, executive, or judicial). Traditionally, lobbyists participate in direct lobbying. A survey of Washington lobbyists showed that 98% use direct lobbying with government officials when promoting their groups' interests (Schlozman & Tierney, 1986). Grassroots lobbying requires lobbying interest group members and others outside the organization. When lobbyists do not make sufficient progress with policymakers, they may ask group members and others to write letters and participate in other methods to attract the interest of policymakers.

Lobbying legislators is the strategy that receives the most attention. If an interest group's goal is new legislation, lobbyists will meet with lawmakers in an effort to promote the interest group agenda. They meet with legislators in an effort to garner support for the group's position and to provide expert information, advice, and language when the lawmakers write proposed legislation. However, most interest groups do not have sufficient money to give to

reelection campaigns and effectively sway a legislator's vote, so they tend to be just providers of information. Lobbyists also participate in coalition building. They see an advantage in having a diverse coalition. If many organizations and associations join together, they can send a strong message to the policymakers.

Interest groups do not always look to make new laws. Often, they lobby for a favorable interpretation of existing regulations. For example, an environmental interest group may lobby the Environmental Protection Agency for stricter interpretation of existing law. Interest groups lobby the executive departments and agencies in much the same ways they lobby legislators, including through personal contacts, offering research, and general public relations.

Although interest groups primarily affect change through lobbying legislators and administrators, sometimes they may need to take their cause to the judiciary. Through class-action lawsuits and *amicus curiae* (friend of the court) briefs, interest groups can affect the judiciary. Education, women's rights, the death penalty and criminal justice issues, abortion, and prayer in public schools are some of the social issues brought to the U.S. Supreme Court by interest groups. The NAACP is a classic example of how an interest group used the courts to affect change. Thurgood Marshall, who later served as an associate justice of the U.S. Supreme Court, served as the chief attorney for the NAACP (an interest group) in its challenge to the constitutionality of state laws requiring racial segregation in public schools (*Brown v. Board of Education*, 1954). After successfully challenging this issue, the NAACP went on to successfully challenge other venues of segregation, such as public transportation, restaurants, hotel accommodations, and others. Like the NAACP, the American Civil Liberties Union (ACLU) has represented the interests of many smaller groups and individuals who otherwise would not have the funds to bring a case to the U.S. Supreme Court. For example, in 1967, the Lovings did not have the funds to fight the antimiscegenation statute of Virginia. The ACLU represented the Lovings' interests, resulting in the overturning of all antimiscegenation laws in the United States (*Loving v. Virginia*, 1967).

Lobbying is the process of petitioning government to influence public policy and is an important part of the democratic process. The right to form groups and to present the groups' ideas is guaranteed by the Bill of Rights. Today, because of the size and complexity of government, individual citizens cannot effectively petition the government for redress of grievances. Federal and state governments deal with large numbers of complex economic and political issues, allowing little time for direct constituency contact. Just as few people would think of appealing to the courts without the benefit of counsel, few people would consider dealing with government without the benefit of a lobbyist. This has given rise to professional lobbyists. A lobbyist is a person who attempts to influence government policies and actions. Interest groups, such as business,

professional, ideological, and trade associations may employ lobbyists.

Since Americans who are disadvantaged economically cannot afford to join interest groups and may not have the time or expertise to find out what group might represent them, some scholars suggest that interest groups and lobbyists are the privilege of the upper and middle class and those who belong to unions and special interests. The poor depend largely on indirect representation. Issue-oriented interests may represent lower-income groups if the policy positions involve the problems of the poor but are not limited to the poor, such as health care, affordable housing, social security, and education. Although they may not join a group, they reap the benefits. Additionally, efforts on behalf of the poor come from public housing officials, welfare workers, public interest groups, and other groups that speak indirectly for the poor. The poor remain outside the interest group network and have little direct voice of their own.

Lobbying is a difficult job, demanding hard work and long hours. It is demanding not simply because it is difficult to get government to do what you want it to do, but because there are usually other lobbyists working on the same issue asking government to do something else. To stand out from the competition, lobbyists must be seen as reliable sources of information. Optimally, they want to be seen as the real policy experts. Lobbyists perform useful functions by providing information to government, educating the public, and preparing legislation. There are those that suggest lobbyists are a third house of Congress. However, there is no doubt that lobbyists are a powerful force in the American political system. Many interest groups have Washington, D.C., offices and employ full-time lobbyists. Others are lawyers, public relations specialists, or Washington insiders and are hired on a fee-for-service basis.

Political Action Committees

Following Watergate in the early 1970s, Congress limited the amount of money that individuals and corporations could contribute directly to candidates, in an effort to limit the influence of big money in politics. But there was no prohibition on individuals and businesses to organizing committees and donating money in ways that favor one candidate over another. This was the beginning of the enormous growth of political action committees (PACs). This rise in PACs illustrates how interest groups were able to work around the reforms so important to Congress. Although PACs grew out of this reform, congressional members benefit from PACs. Generally, the major portion of PAC contributions go to incumbents, effectively securing reelection.

The primary way to gain access to legislators is the campaign contribution. Interest group contributions not only help lobbyists gain access, but also help elect people friendly to the group's goals. As the cost of campaigns continues to rise, legislators must depend more heavily on

the contributions of organized interest groups. PACs are created to raise and distribute funds for campaign and election purposes. Critics of PACs state that the money contributed to finance campaigns leads to influence, and influence in a democracy should not be tied to money.

How Do Interest Groups Form?

The proliferation of interest groups began after World War II, and they continue to grow. As government expands, so do interest groups. President Franklin Roosevelt, through his New Deal policies, expanded the reach of government, effectively creating an atmosphere that has created the enormous number of interest groups we have today. The role of interest groups is to influence government and affect change. As government expands, interest groups grow in an effort to influence a larger government.

In recent decades, the growth of interest groups can be largely attributed to the creation of public interest, citizens' groups, or single-issue groups. Interest groups will form when the political environment supports private interests and when the rules make it easy to organize. Diversity of interests, such as race, religion, and ethnicity, has contributed to the growth of interest groups. As we become a more global society, interest groups will form to accommodate the global interests.

The number of interest groups began to escalate in the late 1960s and has continued to grow steadily. Interest groups seem to flourish in a society with diverse interests, a political culture that supports the pursuit of private interests, rules that make it easy to organize, and government that is sufficiently active for its policies to have consequences for private parties. All of these conditions are present in the United States. Additionally, the First Amendment guarantees citizens basic rights that are essential to the ability of citizens to form organizations.

Political Culture

Almond and Verba (1965) differentiate among parochial, subject, and participant political cultures. In a parochial political culture, citizens have little awareness of or orientation toward either the political system as a whole or the citizen as a political participant. In a subject political culture, such as Germany, citizens are oriented toward the political system, but they are essentially passive. The participant political culture has a high level of political awareness along with significant citizen participation. The United States serves as a good example of participant political culture. Through citizen participation and awareness, individuals and groups can influence decision making. There is an assumption that more demands will be made on government in a participant political culture.

Daniel Elazar (1984) stated that there are three identifiable political cultures. Individualistic political culture

emphasizes private concerns and views government as a utilitarian tool to be used to accomplish what the citizens want. Moralistic political culture views government as a mechanism for advancing the public interest; governmental intervention in the economy is accepted, and there is more public concern about policy issues. The traditionalist political culture takes a paternalistic and elitist view of government and looks to maintain the existing social order; citizens are expected to be relatively inactive in politics.

For interest groups to survive, they must attract paying members. For example, the AARP, the largest membership organization for people 50 years of age and older, continues to grow because of the growing number of people reaching that age. As baby boomers continue to age and join the organization, the AARP will continue to grow. Interest groups such as the AARP attract resources by giving people something of value. In this case, the AARP represents its members in an effort to promote the interests of older Americans. In exchange, the group members give financial support by paying membership fees.

James Q. Wilson (1973) noted three different kinds of benefits. Purposive benefits are those that are ideological in nature. Material benefits are those that promise some kind of financial benefit to participants—for example, the AARP offers health insurance benefits. Solidarity benefits are those that are derived from directly participating with others in a group endeavor—for example, a local chapter of the AARP may ask members to canvas the community to promote the organization's agenda, or they may get together to stuff envelopes.

The Free Rider

Those who get the benefit without having to join the organization are referred to as free riders. Free riders often make it difficult for interest groups to raise the necessary funds to function as needed. For example, public television depends on contributions from viewers to support its programs and budget. However, if one does not donate money to public television, one still gets the benefit of its programming.

If a trade union negotiates a pay increase for industry workers, all employees will benefit even if they are not members of the union. However, there is still a benefit to paying dues to a group. Although all employees may benefit from a pay increase, there are selective benefits for members. Members of a union may enjoy strike benefits, insurance benefits, or the ability to go to union meetings or conventions for professional-growth purposes.

Olson (1965) wrote that if everyone in a group has interests in common, then all will act collectively to achieve them, and in a democracy, the greatest concern is that the majority will tyrannize and exploit the minority. He argued that individuals in any group attempting collective action will have incentives to free ride on the efforts of others if the group is working to provide public

goods. Individuals will not free ride in groups that provide benefits only to active participants.

Many interest groups are ideological in purpose and are not so dependent on membership. In these groups, members drop in and drop out. For example, in recessionary times, if one needed to choose between joining a professional group (trade union, bar association, teachers union, etc.) or an ideological group such as the Sierra Club or a politically motivated organization, it is likely that one would choose a professional group. Citizens are less likely to drop out of professional associations.

The First Amendment right to assembly implies a fundamental right to form organizations. Interest groups coupled with federalism make government more accessible to citizens. Although the Constitution established one government, it recognizes state governments. This is the foundation of federalism. Americans' access to government is easier at the state, county, and local levels. Arguably, federalism fosters individual participation in the political process. Those interested in entering the political arena often start with an interest group at the local level. This environment shapes the political culture.

Conclusion

Interest groups are an integral part of the political landscape of the United States. James Madison (1787) viewed interest groups as a necessary evil. He believed that interest groups not only conflict with each other but also conflict with the common good. However, Madison thought that abolishing interest groups would destroy liberty—a remedy “worse than the disease.”

As stated in a previous section, the socioeconomic status of group membership plays a key role in its ability to influence and promote its agenda. The AARP has a very large membership of educated and committed members. As a result, when the AARP speaks, Congress listens. Those who complain that their interests are not represented, or that one group has too much access to lawmakers, can form groups. As Madison implies, we are best served by having multiple groups, so that one group does not dominate. Nevertheless, socioeconomic status is an important part of interest group influence. “The flaw in the pluralist heaven is that the heavenly chorus sings with a strong upper-class accent” (Schattschneider, 1960, p. 35).

References and Further Readings

- Almond, G. A. (1958). Research note: A comparative study of interest groups and the political process. *American Political Science Review*, 52(1), 270–282.
- Almond, G., & Verba, S. (1965). *The civic culture*. Boston: Little, Brown.
- Berry, J. M. (1996). *Lobbying for the people*. Lincoln: University of Nebraska Press.
- Brown v. Board of Education, 347 U.S. 483 (1954).

- Dahl, R. (1961). *Who governs?* New Haven, CT: Yale University Press.
- Dahl, R. (1989). *Democracy and its critics*. New Haven, CT: Yale University Press.
- Deutsch, K. (1970). *Politics and government*. Boston: Houghton Mifflin.
- Elazar, D. (1984). *American federalism: A view from the state*. New York: Harper & Row.
- Henninger, D. (2010, January 21). The fall of the house of Kennedy. *Wall Street Journal*. Available at <http://online.wsj.com/article/SB10001424052748704320104575015010515688120.html>
- Janda, K., Berry, J., & Goldman, J. (1997). *The challenge of democracy*. Boston: Houghton Mifflin.
- Loving v. Virginia, 388 U.S. 1 (1967).
- Madison, J. (1787). *Federalist papers: Federalist No. 10*. Retrieved February 14, 2010, from http://thomas.loc.gov/home/fedpapers/fed_10.html
- Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Olson, M. (1982). *The rise and decline of nations: Economic growth, stagflation, and social rigidities*. New Haven, CT: Yale University Press.
- Ornstein, N. J., & Elder, S. (1978). *Interest groups, lobbying, and policymaking*. Washington, DC: CQ Press.
- Schattschneider, E. (1960). *The semisovereign people*. Chicago: Holt, Rinehart & Winston.
- Schlozman, K. L., & Tierney, J. T. (1986). *Organized interests and American democracy*. New York: Harper & Row.
- Stern, A. (2010, February 3). The public union ascendancy: Government union members now outnumber private for the first time. *The Wall Street Journal*. Available at <http://online.wsj.com/article/SB10001424052748703837004575013424060649464.html?KEYWORDS=The+Public+Union+Ascendancy%3A+Government+union+members+now+outnumber+private+for>
- Tocqueville, A. de. (1834-1840). *Democracy in America* (H. Reeve, Trans.). London: Saunders and Otley.
- Truman, D. (1951). *The governmental process: Political interests and public opinion*. New York: Knopf.
- Verba, S., Schlozman, K. L., Brady, H., & Nie, N. H. (1993, June). Citizen activity: Who participates? What do they say? *American Political Science Review*, 87, 303-318.
- Wilson, J. Q. (1973). *Political organizations*. Princeton, NJ: Princeton University Press.

88

AMERICAN FEDERALISM

BRIAN P. JANISKEE

California State University, San Bernardino

The United States has a federal political system. Federalism is a system in which political power is shared by national and subnational governments. A national government is the central governing authority in a country. Examples of subnational governments include states, provinces, and territories. Political scientists tend to distinguish federal systems from the two other major categories: confederal and unitary systems. A confederal system is one in which subnational units have nearly all the political power. A unitary system is one in which the national government has nearly all the political power. There are few countries in the world with an identifiable confederal system, with Switzerland being the prominent example. Accordingly, the two main types of political systems in practice throughout the world today are unitary and federal systems.

There are 21 nations that have a federal system of government. These nations account for nearly 40% of the world's population. In addition to the United States, other countries that have federal systems—to one extent or another—include Australia, Belgium, Brazil, Canada, Germany, India, Mexico, Nigeria, and Russia. Although the autonomy enjoyed by subnational units varies from nation to nation, federal systems have a few things in common. Political scientist William Riker (1964) identified three characteristics that federal systems share: (1) National and subnational governments, simultaneously, share power over the same territory and population;

(2) national and subnational governments each have at least one policy area in which they are preeminent; and (3) the autonomy of each level of government is officially recognized. Among democratic nations, the American system is the oldest continuous federal political structure in the world. American federalism divides power between the national government, headquartered in Washington, D.C., and the governments of the 50 states. Additional power is distributed between the national government and federally administered territories, in addition to hundreds of federally recognized Native American tribes.

Models of Federalism in America

From the installation of the Constitution in 1789 until the 1930s, the relationship between the national and state governments could be described as dual federalism (Bowman & Kearney, 2009). Each level of government was understood to be supreme within its own sphere of authority. The national government was supreme in such areas as national defense, foreign relations, mail delivery, and customs enforcement. The state government was supreme in such areas as education and law enforcement. This model allowed for a significant amount of state autonomy with the federal structure.

An extreme version of state-centered federalism is nullification. John Calhoun was the principal advocate of this

idea. This understanding of federalism holds that states—as sovereign political entities—have the authority to ignore, that is, nullify, national laws with which the states may disagree. Although not a predominant view, even among those who favored a state-centered interpretation of federalism, the spirit of nullification contributed greatly to the crisis that eventually led to the Civil War (1861–1865).

Dual federalism continued to be the dominant model of federalism until the 1930s. The inauguration of President Franklin Roosevelt (1933–1945) ushered in the New Deal, a program to bring more national involvement and finances into the provision of services in the states. President Roosevelt succeeded where other national-minded political leaders had failed due, for the most part, to the economic collapse of the Great Depression. Unemployment levels hovered around 20% until the buildup to World War II (1941–1945). Arguments for limited government and state-centered federalism became much less popular during this period. This period of cooperative federalism lasted until the Great Society of Lyndon Johnson.

In the aftermath of President Johnson's (1963–1969) landslide reelection in 1964, the dominance of the national government in the American system of federalism reached new levels. Where the hallmark of intervention of Roosevelt's New Deal was a combination of aid to states and localities, Johnson's Great Society was characterized by direct national aid to local governments. The national government greatly expanded its role in health, education, and welfare policy. These were areas in which the states had traditionally been dominant. This effort, however, was met with some resistance.

With the ascendancy of Richard Nixon (1969–1974) to the presidency, the concept termed *New Federalism* became a rallying point for those who wanted to restore the balance between the national government and the states. Richard Nixon pursued this, in part, through revenue sharing: direct transfers from the national government to states and localities, with the intent of bypassing the federal bureaucracy. The New Federalism program of Ronald Reagan (1981–1989) emphasized the concept of block grants, whereby the national government would send less money to the states in exchange for the states receiving more autonomy in the implementation of programs. This attempt to move power from national to state governments is known as *devolution*.

Despite these attempts to devolve power back to the states, the national government continues as the dominant force in American federalism. The economic crisis that began in late 2008 has further tilted the balance of power in American federalism in favor of the national government. The bank bailout effort during the closing days of the George W. Bush administration and the economic stimulus plan enacted at the start of Barack Obama's administration have ushered in an era of even greater influence by the national government on state policy.

The national government has established itself as the dominant member in the national-state relationship. This is evidenced by the current usage of the term *federal*. At the time of the founding, the term *federal* was synonymous with a system of government that reserved significant authority for state governments. Federal meant decentralized. Presently, the term *federal government* is synonymous with *national government*. When people speak of federal as a political concept, it is synonymous with centralized authority.

Even in this current era of national dominance, policy-makers at all levels of government must confront the complex national-state relationship. Federalism is an important part of the American political tradition. It is, as political theorist Vincent Ostrom (1991) contended, a way of life, not merely a method of administrative organization. The origins of federalism go back beyond the framing of the Constitution to the very establishment of the colonies themselves.

Origins of American Federalism

American federalism has its origins in the method by which the British settled what is now the Atlantic coast of the United States.¹ The settlement of the American Atlantic coast took place by means of colonies. The manner of settlement and political organization of each colony was different. The pace of settlement was uneven. The colonies were formed at different times, under different conditions, and under different monarchs—from 1606 to 1732—under the reigns of King James I, King Charles I, King Charles II, Queen Anne, and King George II.

There were three main categories of colonies: corporate, proprietary, and royal. Corporate colonies were chartered by officially recognized companies. Proprietary colonies were chartered by one person or a small group. Royal colonies were under direct authority of the Crown. Those colonies originally settled as corporate colonies were as follows: Massachusetts (1629, Charles I), Connecticut (1639, Charles I), Rhode Island (1644, Charles I), New Hampshire (1680, Charles II), and Georgia (1732, George II). Those colonies originally settled as proprietary colonies were as follows: Virginia (1606, James I), Maryland (1632, Charles I), Carolina (1663, Charles II), New York (1667, Charles II), Pennsylvania (1681, Charles II), and Delaware (1704, Anne). Those colonies first recognized as royal colonies were as follows: New Jersey (1702, Anne) and the two Carolina colonies—North and South—that were formed when Carolina was split in two (1729, George II). Eventually, the original charters of all the colonies were dissolved by Great Britain, with direct authority reverting to the Crown—except for Connecticut, Maryland, and Rhode Island. Because of these differing experiences—of time, place, and circumstance—each colony had a distinct

identity, a distinctness that was retained in the period after independence. Each colony was different. The different colonies became different states. Therefore, each state was different. According to political scientist Daniel Elazar (1966), the modern differentiation in state policies can be explained, in part, by differences in the political culture of each state.

Any leader who called for greater unity in America would have to confront the existential fact that such a union would have to be formed by entities—states or colonies—that had enjoyed many years as separate political bodies. This tension has been ever present in the American experience. Political historian M. J. C. Vile (1961) understood American federalism as something more than a political system. It is a state of mind that has been engrained in the American tradition for more than 400 years.

Initial Plans for Union

The first significant attempt to unify the colonies under an organized political structure was the Albany Plan of Union in 1754. In this year, war had broken out between France and Great Britain. This war is known as the French and Indian War (1754–1763). French forces in Canada posed a grave threat to colonial security. Meeting in Albany, New York, representatives from the Northern colonies drafted a plan of union for all of British America. The hope was that the proposed union would—among other things—provide for a better defense against a French invasion. Benjamin Franklin—inspired, in part, by the political structure of the Iroquois Confederacy—was the principal architect of this plan. The plan called for a union that would be headed by a president, who would be appointed by the Crown, and a representative body, which would be chosen by the colonial assemblies and proportioned according to the amount each colony contributed to the general treasury. For example, Massachusetts would have been entitled to seven representatives, while Rhode Island would have been entitled to two representatives. Both the Crown and the colonies eventually rejected this plan. Many colonists feared a central authority so close to home. The Crown feared an organized colonial entity that could resist royal authority. The Albany Plan, however, set the precedent for subsequent attempts to form an American union.

The second significant attempt to unify the colonies took place as a result of the War of Independence between the colonies and Great Britain (1775–1783), during which representatives of the colonies—now states—declared their independence from Great Britain. The Declaration of Independence was drafted and signed in 1776. Although the Declaration sets forth principles that could animate the formation of a government, it did not lay out particular arrangements by which the new, independent American political system would be formed. The document that laid out such a system was the Articles of Confederation, which

was drafted in 1777. It was not ratified, however, until 1781, when Maryland finally approved the plan. In the interim between the drafting of the document and the ratification, national affairs were conducted by the Continental Congress. On ratification of the Articles, the Congress of the United States came into being. This Congress, however, was different from the modern version. Each state was entitled to a single vote—regardless of population or wealth. Furthermore, the powers of this new Congress to regulate commerce and to collect taxes to fund such activities as national defense was virtually nonexistent. The Congress had the authority to ask the states for money but had no power to collect. Even though it was poorly funded under such an arrangement, the Continental Army secured victory in the War of Independence. The troubles for America, however, did not dissipate after the official cessation of hostilities in the Treaty of Paris in 1783.

The postwar economy was in a shambles. War debts mounted, and national security threats still loomed on all sides. To address these concerns, representatives from the various states—except Rhode Island—gathered in Philadelphia in the summer of 1787. The initial task of the delegates was to revise the Articles of Confederation so that the country could better deal with the challenges it faced. The gravity of the nation's troubles, however, led many delegates to believe that the Articles of Confederation were beyond repair. New plans for the country's government were introduced.

The Constitutional Convention

There were two main plans under consideration at the Philadelphia convention: the Virginia Plan and the New Jersey Plan. The Virginia Plan was drafted by James Madison. The plan called for the creation of a strong national government. The principal governing body set forth in the Virginia Plan would be a bicameral Congress. The number of representatives each state would have in the Congress would be determined, in effect, by state population. This is a clear rejection of the principle of equal state representation that was embedded in the Articles of Confederation. The New Jersey Plan, proposed by Samuel Patterson, recognized the need for a stronger national government yet maintained the unicameral Congress, with equal representation for each state. As political scientist Martin Diamond (1981) pointed out, the dispute over these plans was not mere bickering between large and small states. Much more was at stake: the legacy of republican theory going back to the ancients.

Although it was widely recognized that the Articles of Confederation were defective, the core dispute was over the appropriate territorial size of a republic. The disagreement over the creation of a national government and the means of representation in the Congress forced the delegates at the convention to confront the very nature of republican government. The popular belief—offered by

the political theorist Montesquieu in *The Spirit of the Laws*—was that republics could function only in smaller territories. After all, ancient republics had predominantly been city-states. According to those inspired by this line of thinking, a republic could not survive in a large nation, and the new constitution, like the Articles of Confederation, must hold dear to the idea of state-centered sovereignty. As scholar Samuel Beer (1993) observed, however, James Madison set about to turn this argument on its head.

Madison's argument in the convention—which later became famous in written form as “Federalist No. 10”—was that republics were safer in large territories. According to Madison, the chief threat to the stability of a republic was an oppressive majority faction—a group driven by a common impulse. In a large territory, however, no single faction could form a majority. There would be too many factions in a nation the size of the United States. This would force factions to form coalitions and, thus, require them to make compromises. Republics in the past had failed because they were small. Although those eager to protect the autonomy of states did not give up on the small republic idea, Madison's argument was persuasive enough for a compromise to take place.

Delegates from Connecticut—Roger Sherman, Oliver Ellsworth, and Samuel Johnson—offered a compromise that eventually broke the deadlock. The Connecticut Compromise—or Great Compromise—was for the creation of a bicameral Congress. The House of Representatives would be composed of representatives selected by the voters in the individual states, with the number of representatives from each state determined by population. This was similar to the method proposed by Madison in the Virginia Plan. The Senate would consist of representatives chosen by state legislatures, with each state receiving equal representation—two per state—that was similar to the method proposed in the New Jersey Plan. Through this compromise, a strong national government could be created while protecting state autonomy. The state governments themselves would provide representatives to one of the two branches of the national legislature.

By September, a draft of the Constitution was ready to be sent to the states for ratification. Many of the same arguments took place in state ratification conventions that took place in the Philadelphia convention. Along with Alexander Hamilton and John Jay, James Madison contributed to a series of articles that became known as *The Federalist Papers*. Writing under the collective pseudonym Publius, these articles were the most systematic and coherent defense of the new Constitution. These writings were used by pro-Constitution delegates at ratification conventions throughout the country. As for the so-called anti-Federalists, writers using pen names such as Cato and the American Farmer argued against the ratification of the Constitution. This collection has come to be known as the *Anti Federalist Papers*. Despite the best efforts of the anti-Federalists, the proposed Constitution was ratified in the late spring of 1788. The

new national–state experiment, which we know today as federalism, would be implemented in 1789.

Federalism in the Constitution

The Constitution, as it went into effect in 1789, contained seven articles:

- I. Legislature
- II. Executive
- III. Judiciary
- IV. National State Relations
- V. Method of Amending Constitution
- VI. Constitutional Supremacy
- VII. Method of Ratification

Article IV specifically deals with national–state issues, though the principles of federalism can be found throughout the Constitution. The document is infused with mention of the need of both states and the national government to share authority and responsibility for the operation of government. By its design, the national and state governments interact throughout the system of government in the United States.

Preamble

The intent of the framers of the Constitution is clear from the start of the document:

We the people of the United States, in order to form a *more perfect union* [emphasis added], establish justice, insure domestic tranquility, provide for the common defense, promote the general welfare, and secure the blessings of liberty to ourselves and our posterity, do ordain and establish this Constitution for the United States of America. (U.S. Constitution, 1787, Preamble)

The government that was to be created by the Constitution was to provide for “a more perfect union.” They did not seek to provide a perfect government system. On the other hand, they did not merely tinker with the Articles of Confederation. This union was designed to bring together the various states under a common government yet allow the states to retain power and identity. However, as discussed in previous sections, there has been controversy—in disputes over nullification, dual federalism, and cooperative federalism—over how to balance the power and identity of the union between state and national governments.

Article I

The first article of the Constitution has a dual character. In one sense, it is the section of the document in which the powers of the legislature are outlined. In a more fundamental sense, however, it is also the section of

the Constitution in which the general powers of the national government are laid out.

*Congressional Elections,
Apportionment, and Representation*

One of the first topics mentioned in Article I was the manner of elections to the bicameral Congress, which would consist of the House of Representatives and the Senate. The House would be selected by the “people of the various states.” If a representative is no longer able to fill the seat—because of death, illness, or resignation—the governor of the state shall call for a special election to fill the seat. Those eligible to vote in state legislative elections would be eligible to vote in House elections. The number of representatives of each state would correspond to the population of the state. Each state would have at least one representative, with the remaining proportion being determined by population.

At the constitutional convention in 1787, the Southern states wanted slaves to be counted for the purposes of representation in the House, even though they were denied basic political rights. It was in the interest of slave-holding states to have slaves count fully in the population, even though they would be denied the right to vote. This would swell the number of House members from Southern states and help these states maintain their power. If the goal were to weaken the power of Southern states in the national government—thus leading to the eventual abolition of slavery—the practical result would have been for slaves to not count at all toward the representation of Southern states in the House. This would leave the Northern states in a clearly dominant position and, perhaps, able to legislate slavery out of existence. The Southern states, however, would have balked at this notion and, most likely, walked out of the convention, leaving the fragile union in tatters. An unsavory compromise was reached in which slaves would count as three fifths of a person for the purposes of representation in the House. This preserved the Union and the hope that, one day, under a strong Union, slavery would be abolished. The alternative would have left the country divided, with the Southern slave-holding states left to form their own union, a union under which slavery could have expanded unabated by any political influence from the Northern states. Martin Diamond (1981) argued, “Not striking the bargain would have freed not a single slave while it would have destroyed the possibility of union” (p. 39). This possibility, with respect to slavery, was finally realized in the aftermath of a brutal Civil War (1861–1865).

In the Senate, the proportion of members from each state was less controversial, at least during the framing of the Constitution.² There would be two members per state, regardless of state population. State legislators themselves would decide who represented their state in the U.S. Senate. The state governments would have their direct

representatives fill the seats of the U.S. Senate, thus providing a valuable check—from the perspective of the states—against potential power grabs from the national government. State governors would appoint replacements in the event of an unexpected vacancy in a Senate seat.

As for the “time, places, and manner” of congressional elections, the Constitution prescribed that the national government would be allowed to determine the time and manner, but the states were reserved the power to determine the places. Furthermore, the national government reserved to itself the power to decide who was duly elected to the House and Senate. The final arbiter of disputed elections would be the Congress, not the states. In addition, members of Congress would be paid out of the national treasury and not the state treasury.

Enumerated Powers

The powers of the Congress and, by extension, the national government are described in Article I, Section 8. The national government would have, among others, the power to collect taxes, borrow money, regulate commerce, administer immigration, coin money, establish a post office, declare war, raise armies, provide a navy, and make laws “which shall be necessary and proper for carrying into execution the foregoing powers” (U.S. Constitution, 1787, Article I, Section 8). These powers have been described by scholars as enumerated powers. These powers are those that are directly given to the national government and listed in a systematic fashion.

There are two main schools of thought on how to interpret the list of powers in Article I, Section 8. At first glance, they appear to be a simple, straightforward accounting of the basic powers any government should have in order to function properly. A strict constructionist reads the list as a limitation on the power of the national government; the national government’s activities are limited to those on the list. A loose constructionist reads other powers—implied powers—into the list that the national government could exercise. For example, the responsibility to regulate commerce may give the government power over a wide variety of activities that are related to or affect commerce. In addition to the commerce clause, the necessary and proper clause has been interpreted by loose constructionists in a manner that would facilitate activity of the national government that goes well beyond those specifically described on the list. For example, the commerce clause could allow the government to provide health insurance, income subsidies, and regulate labor laws.

As for the necessary and proper clause, this could allow the government to establish a national bank, for example, even though the establishment of such a bank is not listed as one of the national government’s delegated powers. In fact, dispute over the constitutionality of a national bank was perhaps—after slavery—the most controversial political

debate in America, from the founding period until the Civil War. The meaning and interpretation of Article I, Section 8 is fraught with both partisan debate and thoughtful deliberation on both sides of the issue.

Prohibitions on State Action

The concluding section of Article I prohibits states from engaging in a wide variety of activities, which are better suited for a national-level government. States are prohibited from entering into treaties or alliances with foreign governments. State governments may not coin money. They cannot, without the permission of Congress, impose import and export taxes on goods passing through their jurisdictions. In addition, states cannot pass laws that impair “the obligation of contracts.” This issue could come to the fore if, for example, a state attempted to provide debt relief to some of its residents by nullifying loans.

Article II

The second article of the Constitution deals with the executive branch, which is headed by the president. The Constitution does not prescribe a national method by which the president is chosen. Instead, the Constitution sets forth a method by which each state appoints electors to perform this task. This body of electors is known as the Electoral College. The number of electors each state may appoint to this body is equal to the number of members each state has in the U.S. House and U.S. Senate. Current members of Congress, however, along with national government officials, are prohibited from serving as electors.

It was assumed by the framers that electors would tend to favor candidates from their own state. Accordingly, electors were required to place two votes for president, with the stipulation that both votes could not be for a candidate from the same state as the elector. The candidate receiving a majority would be president. The candidate receiving the second-highest number of votes would be vice president. This was changed somewhat by the Twelfth Amendment (1804). Electors are now required to make one choice for president and one choice for vice president, with the stipulation that an elector cannot vote for both a president and vice president from the same state if the elector is also from that state.

As for the manner of selecting the electors themselves, it is left to each state legislature to make this decision. Although popular election in each state is the current method by which electors are chosen, this has not always been the case and could change in the future. In fact, in the midst of the recount controversy in Florida during the 2000 presidential election, the legislature of that state appeared willing to select Florida’s electors, if the courts failed to resolve the controversy in a timely fashion. Furthermore, several state legislatures have expressed

interest in awarding their states’ electors to the winner of the national popular vote, regardless of the winner in their own state.

Federalism also appears in the manner of selecting a president if the election aforementioned procedures fail to produce a majority. The House of Representatives, in the original Constitution, would choose the president from among the top five candidates. However, the House members would not vote as individuals. Instead, they would be required to caucus by state, with the state delegation voting as a unit. The winning candidate would have to receive a majority of state delegations to secure the presidency.

In addition to presidential elections, federalism appears in the Constitution’s description of the president’s role as commander in chief and principal negotiator with foreign powers. The president, in times of emergency, may assume command of the various militias in each state. As for foreign negotiations, treaties brokered by a president may not go into effect unless a two-thirds majority of senators approves. Since the Senate is apportioned equally by state, this procedure protects the interests of a minority of states against treaties—such as commercial or trade agreements—that would have a disproportionate effect on such states.

Article III

The third article of the Constitution outlines the judicial power of the national government. Although most court cases are heard at the state level, the Constitution, as drafted in 1787, granted original jurisdiction to the Supreme Court in cases “in which a state shall be a party” (U.S. Constitution, 1787, Article III). The role of the national courts in state matters was further refined in the Eleventh Amendment in 1795.

Article IV

Although the principle of federalism is expressed throughout the Constitution’s seven articles, the fourth article of the Constitution could be named *the Federalism article*. The entire contents are taken up with interstate issues and the relationship between the state and national government. The Constitution requires the “full faith and credit” be rendered to the public acts of one state in that of another. For example, marriages in one state are generally recognized in others. However, the early-21st-century controversy over gay marriage has caused considerable debate over the proper interpretation of this element of the Constitution. In addition to the full-faith-and-credit clause, Article IV also establishes the principle by which the “citizens of each state shall be entitled to all Privileges and Immunities of Citizens in the several States” (U.S. Constitution, 1787, Article IV). For example, this clause guarantees a right to travel from state to state. In addition to this interstate provision, a person charged with a crime

who subsequently flees to another state must be extradited back to the state in which the alleged crime took place.

As for the creation of new states, Congress is given this authority. However, Congress may not create a new state out of part of an existing state or by combining parts of more than one state, without the approval of the affected states. This is a critical protection for states in a federal system. Without this protection, a national government, displeased with a state, may seek to destroy it as a political entity by carving it up into new states or adjoining it to another.

Article V

The fifth article of the Constitution is taken up with the method by which the Constitution is to be amended. The manner in which this is prescribed in the Constitution protects state governments, to some extent, against unilateral action by the national government. The Constitution cannot be amended without the cooperation of state governments. In fact, there is one method by which the states, without the involvement of the national government, could amend the Constitution.

Of the two methods available to amend the Constitution, only one has been employed to date. This method begins with a proposal introduced into the United States Congress. This proposed amendment must be passed by a two-thirds vote of each chamber, voting separately. If the proposed amendment passes, then it is sent to the states. If three fourths of the state legislatures—or state conventions, depending on the nature of the proposal—approve the amendment, it is ratified and becomes part of the Constitution. This last step, however, is even more difficult than it appears. Since 49 of the 50 legislatures are bicameral—with Nebraska as the exception—a proposed amendment must win approval in two legislative chambers in each state in order to have that state count toward the three-fourths requirement.

The second method by which amendments could be added to the Constitution is initiated by the state legislatures. If two thirds of the state legislatures approve an application for a constitutional convention, delegates would gather for the purpose of proposing amendments to the Constitution. These proposed amendments would then be sent back to the states for approval, with three fourths being necessary for ratification, either by legislatures or conventions, whichever is prescribed in the amendment proposal.

Article VI

The sixth article of the Constitution contains the supremacy clause. This clause states that the Constitution is the supreme law of the land. This clause is one of the main bulwarks of the school of thought that holds that the national government should hold the balance of power in the

national–state relationship. In addition to the supremacy clause, the sixth article mandates that both national and state officials “shall be bound by Oath or Affirmation, to support this Constitution” (U.S. Constitution, 1787, Article VI).

Article VII

The seventh article of the Constitution established the method by which the Constitution would be ratified. The document, as drafted in Philadelphia in 1787, was sent to ratification conventions in the 13 states. To be implemented, 9 of the 13 states would need to approve. In effect, a much larger majority would need to be secured. It is doubtful that the new union would have survived if 9 of those states did not include the likes of Virginia, New York, or Pennsylvania. As it happened, the Constitution went into effect in 1789 with 2 states still left to ratify: North Carolina and Rhode Island. North Carolina ratified the Constitution in November of that same year, with Rhode Island delaying ratification until 1790.

Amendments

There have been 27 amendments to the Constitution. Those with a significant impact on the operation of federalism include the following: the Sixth, Tenth, Eleventh, Thirteenth, Fourteenth, Fifteenth, Sixteenth, Seventeenth, Twenty-Fourth, and Twenty-Sixth Amendments. The Sixth Amendment, ratified in 1791 as part of the Bill of Rights, established the right of the accused to a speedy trial by an impartial jury in the state in which the crime was committed.

The Tenth Amendment assigns those powers not specifically delegated to the national government to be reserved for the states or the people. Just as those who advocate a national-centered view of federalism point to the supremacy clause in Article VI of the Constitution, those who advocate a state-centered view of federalism point to the language of the Tenth Amendment. This amendment was also ratified in 1791 as part of the Bill of Rights. The Eleventh Amendment, ratified in 1795 in response to the Supreme Court’s decision in *Chisolm v. Georgia*, held that the Supreme Court could hear cases between private citizens and states. The Eleventh Amendment grants states protection against federal lawsuits from citizens of other states and from citizens of other countries.

The Thirteenth, Fourteenth, and Fifteenth Amendments—ratified in 1865, 1868, and 1870, respectively—deal with the aftermath of the Civil War. The Civil War was the most contentious and most violent conflict in the American history of federalism. Eleven Southern slave-owning states seceded from the union, on the election of Abraham Lincoln to the presidency. This crisis manifested itself, primarily, in a dispute over the future expansion of slavery into the Western territories. In the course of the war, the elimination of slavery became the goal of the union. On

the war's conclusion, these three amendments abolished slavery (Thirteenth), mandated equal protection under the laws (Fourteenth), and established voting rights for former slaves (Fifteenth).

The Sixteenth and Seventeenth Amendments, both ratified in 1913, contributed greatly to the subsequent dominance of the national government over the states. The Sixteenth Amendment established the constitutionality of a national income tax. This provided the federal government with the opportunity to establish a consistent and significant source of revenue, from which the potential could be realized to fund larger national government programs. The usage of such funds to influence state policy is known as fiscal federalism.

The Seventeenth Amendment changed the method by which senators were elected. In the original Constitution, senators were selected by state legislatures—a vital element of the national–state balance of power. With the ratification of the Seventeenth Amendment, senators were directly elected by the voters. State governments would no longer have a direct hand in the selection of legislators. According to political scientist Ralph Rossum (2001), the balance between national and state government, arguably, has not been the same since.

The Twenty-Fourth and Twenty-Sixth Amendments relate to the right to vote. The Twenty-Fourth Amendment was ratified in 1964 to eliminate the poll tax in elections for national office. Some Southern states used this practice to prevent African Americans from exercising their right to vote. This issue, along with desegregation, was a legacy of the conflict in the Civil War. The Twenty-Sixth Amendment was ratified in 1971 and established a uniform, nationwide voting age of 18 years. Until then, some states still had a minimum voting age of 21. The draft policy in the Vietnam War—with a minimum-age requirement of 18—created political momentum for this amendment.

Supreme Court Cases

It is a characteristic of the American system that policy debates become constitutionalized. This is no less true of controversies over the proper balance of state and national powers within the American system of federalism. As is the case with many areas of policy, Supreme Court cases have proved influential in how the balance of power is decided between the national and state governments. A sample of influential cases follows. Although not exhaustive, this list is a representative sample of the kind of issues that can arise out of the complexities of the American system of federalism:

- *Chisolm v. Georgia* (1793): The Court ruled that the Supreme Court could hear disputes between citizens and states. This was overruled by Eleventh Amendment.
- *Fletcher v. Peck* (1810): The Court ruled a state law unconstitutional.

- *Martin v. Hunter's Lessee* (1816): This established supremacy of the Supreme Court over state courts in national matters.
- *McCulloch v. Maryland* (1819): This expanded the powers of the national government through the necessary and proper clause.
- *Gibbons v. Ogden* (1824): This affirmed the national role in the regulation of interstate commerce.
- *Worcester v. Georgia* (1832): This recognized tribal sovereignty with respect to the states.
- *Cooley v. Board of Wardens* (1851): The states can regulate interstate commerce in limited instances.
- *Wabash v. Illinois* (1886): The ability of states to regulate interstate commerce is limited.
- *NRLB v. Jones* (1937): This significantly enhanced the power of the national government through a broad interpretation of the commerce clause.
- *United States v. Darby* (1941): This affirmed the power of the national government to regulate wages and working hours through the commerce clause.
- *Wickard v. Filburn* (1942): This affirmed the power of the national government to regulate prices through the commerce clause.
- *Cooper v. Aaron* (1958): State law may not contradict Supreme Court decisions.
- *National League of Cities v. Usery* (1976): The power of the national government to regulate commerce has limits under the Tenth Amendment.
- *Garcia v. San Antonio* (1985): The power of the national government to exercise authority through the commerce clause is limited only by the intentions of Congress.
- *South Dakota v. Dole* (1987): The national government can impose drinking age standards on states as a condition of federal funds.
- *U.S. Term limits v. Thornton* (1995): The states cannot set term limits on members of Congress.
- *United States v. Lopez* (1995): This set limits on the national government's power through the commerce clause.
- *Seminole Tribe of Florida v. Florida* (1996): This affirmed the sovereign immunity of states under the Eleventh Amendment.
- *Printz v. United States* (1997): This provided some protection for state government officials against being forced to administer federal law.
- *United States v. Morrison* (2000): This further expanded on the Court's decision in *United States v. Lopez*.
- *Gonzalez v. Raich* (2005): National power under commerce clause extends to regulating use of medical marijuana, even if state law is to the contrary.

Conclusion

The United States is one of 21 nations that have a federal political system. Federalism is a model of government that distributes political power between a national government

and subnational governments. Political power in the United States is balanced between the national government, headquartered in Washington, D.C., and each of the 50 state governments. Political power is also dispersed between the national government and U.S. territories, along with officially recognized Native American tribes. The American system of federalism has its origins in the method of European settlement on the North American continent. The British colonies were formed at different times, with different charters, under different conditions. This contributed to a diversity of state experiences that made a balance of power between national and state governments a practical necessity at the nation's founding. From 1789 until the New Deal of the 1930s, national and state government operated under the principle of dual federalism—each remaining supreme within its traditional policy areas. With the onset of the New Deal, the national government became more involved in policy areas that had customarily been reserved to the states. By the 1960s, the national government had been firmly entrenched as the dominant power in the national–state balance. This is symbolized by the current usage of the term *federal*. At the time of the founding, the term *federal* was synonymous with *decentralized*. Today, the term *federal* is synonymous with *national* and connotes a centralized approach to governing. Even though the national government is dominant, federalism is an important tradition in the American polity and, as political scientist Thomas Dye (1990) is keen to note, “a defense against tyranny” (p. 5). Disputes over the proper balance of powers in the national–state relationship are common, and many of these disputes have been settled by Supreme Court decisions.

Notes

1. Until the Act of Union in 1707, which united England and Scotland under a single monarch, the term *English colonies* may be more accurate. However, in the interest of simplicity, *Great Britain* or *British* will be used in all cases.

2. The controversy over the North South balance in the Senate took place when territories petitioned to join the Union as states. The Missouri Compromise of 1820, The Compromise of 1850, and the Kansas Nebraska Act of 1854 were central disputes over the admission of slave and free states to the Union.

References and Further Readings

- Beer, S. H. (1993). *To make a nation: The rediscovery of federalism*. Cambridge, MA: Harvard University Press.
- Bowman, A. O., & Kearney, R. C. (2009). *State and local government: The essentials*. Boston: Houghton Mifflin.
- Bryce, J. (1911). *The American commonwealth* (Vol. 1). New York: Macmillan.
- Calhoun, J. C. (2002). *A disquisition on government; and a discourse on the Constitution and government of the United States* (R. K. Cralle, Ed.). Union, NJ: Lawbook Exchange.

- Croly, H. (1965). *The promise of American life*. Indianapolis, IN: Bobbs Merrill.
- Diamond, M. (1981). *The founding of the democratic republic*. Itasca, IL: F. E. Peacock.
- Dinan, J. (2008). The state of American federalism 2007–2008: Resurgent state influence in the national policy process and continued state policy innovation. *Publius*, 38, 381–415.
- Dye, T. R. (1990). *American federalism: Competition among governments*. Lexington, MA: Lexington Books.
- Elazar, D. J. (1966). *American federalism: A view from the states*. New York: Thomas Y. Crowell.
- Friedrich, C. J. (1968). *Trends in federalism in theory and practice*. New York: Praeger.
- Goldwin, R. A. (Ed.). (1971). *A nation of states* (2nd ed.). Chicago: Rand McNally.
- Jaffa, H. V. (2000). *A new birth of freedom: Abraham Lincoln and the coming of the civil war*. Lanham, MD: Rowman & Littlefield.
- Jaffa, H. V. (2009). *Crisis of the house divided: An interpretation of the issues in the Lincoln Douglas debates*. Chicago: University of Chicago Press.
- Kesler, C. R. (2003). Introduction. In C. Rossiter (Ed.), *The federalist papers* (pp. vii–xxxi). New York: Signet Classic.
- Kincaid, J., & Tarr, G. A. (Eds.). (2005). *Constitutional origins, structure, and change in federal countries*. Montreal, QC, Canada: McGill Queens University Press.
- Osgood, H. L. (1957). *The American colonies in the seventeenth century* (Vols. 1–3). New York: Columbia University Press.
- Osgood, H. L. (1958). *The American colonies in the eighteenth century* (Vol. 1–4). New York: Columbia University Press.
- Ostrom, V. (1991). *Meaning of American federalism*. San Francisco: Institute for Contemporary Studies Press.
- Patterson, J. T. (1969). *The New Deal and the states: Federalism in transition*. Princeton, NJ: Princeton University Press.
- Peterson, P. E., Rage, B. G., & Wong, K. K. (1986). *When federalism works*. Washington, DC: Brookings Institution.
- Purcell, E. A., Jr. (2007). *Originalism, federalism, and the American constitutional enterprise: A historical inquiry*. New Haven, CT: Yale University Press.
- Riker, W. H. (1964). *Federalism: Origin, operation, significance*. Boston: Little, Brown.
- Riker, W. H. (1987). *The development of American federalism*. Boston: Kluwer Academic.
- Rossum, R. A. (2001). *Federalism, the Supreme Court, and the Seventeenth Amendment: The irony of constitutional democracy*. Lexington, MA: Lexington Books.
- Tocqueville, A. de. (2000). *Democracy in America* (H. C. Mansfield & D. Winthrop, Eds.). Chicago: University of Chicago Press.
- U.S. Constitution. (1787). Available at <http://www.usconstitution.net/const.html>
- Vile, M. J. C. (1961). *The structure of American federalism*. Oxford, UK: Oxford University Press.
- Walker, D. B. (2000). *The rebirth of federalism: Slouching toward Washington*. Chappaqua, NY: Chatham House.
- Weissert, C. S., & Scheller, D. (2008). Learning from the states? Federalism and national health policy. *Public Administration Review*, 68, 162–173.
- Zimmerman, J. F. (2008). *Contemporary American federalism: The growth of national power*. Albany: SUNY Press.

89

AMERICAN POLITICAL PARTIES

JEFFREY KRAUS

Wagner College

Although not provided for in the Constitution, political parties emerged in the years immediately following the establishment of the Republic. Initially, they were loose factions of officeholders, what Duverger (1964) would classify as “cadre parties,” and the development of political parties corresponded to the expansion of the franchise, as political parties evolved into mass parties that focused on organizing and mobilizing the expanded electorate. By the late 19th century, urban political party organizations, known as *political machines* and led by bosses, mobilized immigrant voters (by offering patronage and petty favors) and triggered a reform movement that aimed to reduce the influence of the bosses and party organizations. During the 20th century, as the welfare state emerged and as candidates turned to candidate-centered organizations to run their campaigns and consultants to manage them, the American political party has continued to evolve.

In examining the modern American political party, it is important to understand what they are, their functions and components, and how the present two-party system has developed. However, one cannot ignore third parties in American politics, and many political scientists continue to speculate about the future of parties in the American polity and whether it's time for a new party—or perhaps a new party system.

American Political Parties Defined

Robert Huckshorn (1984) defined the political party as “an autonomous group of citizens having the purpose of making nominations and contesting elections in hope of gaining control over governmental power through the capture of public offices and the organization of the government” (p. 10).

According to John K. White (2009), two contemporary models of political parties have been offered by scholars: the rational-efficient model and the responsible parties model. The rational-efficient model was first presented by Anthony Downs (1957), who contended that winning elections is the principle purpose of political parties. Downs contended that politicians are interested primarily in securing the perquisites of power: “Parties formulate policies in order to win elections, rather than win elections in order to formulate policies” (p. 28). Downs asserted that voters act rationally by using information provided by candidates and parties to make choices that they believe will improve their economic or physical security.

In 1950, the American Political Science Association's Committee on Political Parties (1950) released its report, “Toward a More Responsible Two Party System.” The Committee argued that parties should develop programs and then carry out those programs when their candidates are elected to office. Such parties would offer voters clear

choices and give the victorious party a mandate to govern, making them responsible to the electorate for their actions. Voters could then, in the next election, retrospectively vote on the outcomes produced by the party in power.

The Functions of Political Parties

Political parties perform a number of functions for the American political system. First, they provide symbols for partisan identification, which provides citizens with a basis for participation in politics (Rosenblum, 2008). Second, parties help socialize and educate voters by making them aware of the issues and by encouraging their participation within the established political processes, playing an important role in channeling social conflict. A third function of parties is the recruitment and nomination of political candidates; what distinguishes parties is this function of nominating candidates. By sponsoring candidates for public office, parties provide a form of quality control. As Janda, Berry, and Goldman (2009) noted, “Party insiders, the nominees’ peers, usually know the strengths and faults of potential candidates much better than average voters do and thus can judge their suitability for representing the party” (p. 231). Once they nominate candidates, parties mobilize voters to support those candidates. Parties also present proposals to voters during election campaigns and help facilitate cooperation between the members of the party in government. Our system of separation of powers within a federal state divides power; the parties, through the cooperation of party members in different branches of government and at different levels of government, bring some cohesion to the processes of governing.

The Components of the Party

Political scientists suggest that there are three distinct elements to the American political party: the party in government, the party in the electorate, and the party organization.

The term *party in government* refers to all of the elected and appointed officials who identify with a political party. The members of the party in government work to carry out proposals set forth in party platforms or presented in campaigns or developed by party members in government. It is expected that the members of the party in the different political institutions and at different levels of government will coordinate their activities to ensure enactment of their party’s proposals that will result in continued electoral success for the party.

This coordination of activity to pursue a partisan political agenda was recently on display during George W. Bush’s presidency (2001–2008). Observing this development, Cigler and Loomis (2006) wrote that “with President Bush’s polarizing leadership style and House Majority Whip Tom DeLay’s powerful mix of incentives, cash, and

sanctions, national Republicans have sought to govern through a tight-knit majority that leaves little room for representation of minority party views” (p. 3). DeLay’s K Street Project coerced interest groups and lobbyists into supporting Republican candidates and policies and hiring party operatives for positions with their firms and organizations. Barbara Sinclair wrote that “interest groups are being forced to align with one party or the other, to become part of one of two durable coalitions, and this is true of even groups that would prefer to play both sides of the fence” (cited in Cigler & Loomis, pp. 5–6).

Party in the electorate is a term political scientists use to refer to voters who identify with a political party and who usually vote for candidates nominated by that party. In other political systems, individuals may actually buy party memberships or must be reviewed by a party membership committee before being allowed to join. Membership allows voters to participate in many party activities, notably voting for the party’s candidate for legislative office. In the United States, although voters may join political organizations associated with political parties (including chapters of the Young Democrats or College Republicans), this is not required. Party identification is a psychological attachment to a party, which often shapes the voters’ attitudes about the issues of the day and is a significant determinant in how they cast their votes. In states that hold closed primaries to select candidates, voters may enroll in a political party that will permit them to vote in that party’s primary. However, the party has no control over who enrolls in the party. During the 2008 presidential primaries, conservative talk-show host Rush Limbaugh launched what he called Operation Chaos, urging Republican voters to switch parties to create a muddled Democratic Party contest.

The term *party identification* was first used by Campbell, Gurin, and Miller (1954) in *The Voter Decides*, where they defined party identification as “the sense of personal attachment which the individual feels toward the group of his choice” (p. 89). They suggested that party identification was passed down from parent to child as part of the political socialization process. Party identification was refined by Campbell, Converse, Miller, and Stokes (1960) in *The American Voter*, where they described it as “the individual’s affective orientation to an important group-object in his environment” (p. 121). They contended that most voters based their votes on party identification.

In American politics, voters are identified as Republicans, Democrats, or Independents (voters who state that they do not identify with one of the two major parties). Although a majority of voters still identify with one of the two major parties, the number of voters who identify themselves as Independents has increased. A study by the Pew Research Center for People and the Press (2009) found that 39% of the electorate identified themselves as Independents, a figure that “equals its highest levels in 70 years” (p. 1). The same survey found

that 33% of the electorate identified themselves as Democrats and 22% as Republicans.

The third component of the political party is the party organization. The party organization is composed of the permanent structures and procedures that maintain the party on a daily basis. American political party organizations are distinctive in that they are decentralized, reflecting the federal system. At the national level, the Democrats and Republicans have similar structures. Every 4 years, each party holds a national convention, where they formally nominate their presidential and vice presidential candidates, adopt a platform, and officially kick off their general election campaign. The Democratic Party's (2007) charter states that "the National Convention shall be the highest authority of the Democratic Party" (Article Two, Section 2).

The ongoing functions of the national parties are carried out by their national committees. The national committees consist of representatives from each state party organization; the Democratic National Committee (DNC) also includes representatives from various groups of Democratic elected and appointed officials, as well as constituencies that include Young Democrats, College Democrats, women, and senior citizens. Historically, the national party committees organized the conventions and the general election campaigns for their national tickets. In recent years, the national campaigns have been mounted by candidate-centered organizations (also known as principal campaign committees) while the national committees have broadened their roles. Under Ray Bliss (1965–1969), the Republican National Committee (RNC) initiated a number of reforms designed to make it a more professional organization that could recruit candidates and political operatives and provide an array of services to candidates and state party organizations (Conley, 2008). Today, both national committees are engaged in fund-raising, candidate and campaign staff training, polling, and research. As DNC chair, Howard Dean (2005–2009) attempted to rebuild the Democratic Party nationally through a 50-state strategy, where the DNC deployed staff and resources in traditionally Republican states in an effort to expand the party's reach. Although Dean was initially criticized by party notables, such as political consultants James Carville and Paul Begala and former President Bill Clinton, others suggest that it was the blueprint for Barack Obama's winning 2008 campaign strategy (Berman, 2008).

Each national committee elects a chair. The president usually selects the chair of their party while the party out of power often holds a competitive election. In 2010, the national committee chairs were Tim Kaine (Democrat), the former governor of Virginia (2006–2010), and Michael Steele (Republican), the ex-lieutenant governor of Maryland (2003–2007) and candidate for the United States Senate (2006).

Each of the major parties has state party organizations in each of the 50 states and the District of Columbia. These

organizations are governed by state election laws and their own bylaws. Typically, there is a state chair who is elected by the party's state committee, which is made up of representatives from political subdivisions within the state. For example, the Democratic and Republican State Committees in New York are each composed of a male and a female representative from each of the 150 state assembly districts. In Texas, each of the state committees is composed of a chair, vice chair, and two members from each of the state's 31 state senate districts. In Maryland, the party state committees are made up of representatives from the county committees and the Baltimore city committee.

State political party organizations have a number of functions: candidate recruitment, fund-raising, development of a platform for statewide elections, recommending applicants for state boards and commissions, voter mobilization, and selection of the party's slate of candidates for the electoral college. Although most states provide for the selection of candidates through primary elections, some states grant the parties some latitude. The Virginia code provides that "the duly constituted authorities of the state political party shall have the right to determine the method by which a party nomination for a member of the United States Senate or for any statewide office shall be made" (*West's Annotated Code of Virginia*, 2003, Section 24.2–509A). It should be noted that party organizational strength in each state varies, depending on the political environment within that state (Morehouse & Jewell, 2003), which includes the nature of competition between the two parties within the state (Rosenthal, 1995).

One of the more recent organizational developments at the state level has been the emergence of state legislative campaign committees (LCCs; Shea, 1996). Patterned after the congressional campaign committees, these LCCs have become what Shea called "the '800 pound gorillas' of legislative politics" (p. 11). Shea, a former operative with the New York State Democratic Assembly Campaign Committee, explained that LCCs do the following:

give money, but more important they provide expertise and assistance. A candidate can expect to be invited to Washington, D.C., or a state capital to attend a training session. They can use the media studios to make radio and television spots, and they receive assistance on their direct mail. In some states, LCC operatives join races in the field, essentially running the show. Perhaps the foremost help provided by these units is what Herrnson has termed a "brokerage role." That is, they link candidates with potential contributors and discount service vendors. If a candidate is interested in a survey, for example, they can be put in contact with a top notch pollster and receive this service at a bargain price. . . . They also bring PAC decision makers and candidates together as part of a fund raising match making service. (pp. 11–12)

Political parties are also organized at the grassroots level, with local committees mobilizing voters to support

their candidates. During the latter part of the 19th century and the first half of the 20th century, politics in many urban areas was dominated by so-called political machines. These machines controlled city governments (and had influence over statewide elections) through their ability to mobilize immigrant and working-class voters. These organizations were hierarchical in nature and usually controlled by a so-called boss, who maintained power by dispensing patronage to supporters and the granting of municipal franchises and contracts to businessmen who were willing to kick back some of their profits to the organization. Perhaps the most notorious machine was Tammany Hall, the Democratic Party organization that was a major force in New York City politics for more than a century. William Marcy Tweed, who led the Hall from 1858 to 1871, epitomized the corrupt boss. Between 1865 and 1871, Tweed and his ring stole at least \$50 million from the city (Allen, 1993).

During the Progressive Era (1890–1920), there were a number of proposals intended to undermine the bosses. These included the selection of candidates through primaries to reduce the ability of the bosses to pick candidates, the introduction of the merit system in public employment to reduce the patronage controlled by party leaders, and changes in local governance (the replacement of ward-based city councils with at-large elections) and elections (non-partisan elections) that would reduce the power of the parties. While some machines persisted into the 1970s, notably the Cook County (Illinois) machine led by Chicago's Mayor Richard Daley, most of these organizations lost influence as their traditional supporters moved to the suburbs and the percentage of the vote being cast in urban areas declined (Schneider, 1992).

A Two-Party System

George Washington (1940), the first president of the United States, was “above party” and warned against political parties in his 1796 farewell address, writing that “the common and continual mischiefs of the spirit of party are sufficient to make it the interest and duty of a wise people to discourage and restrain it” (p. 227). Notwithstanding Washington's admonition, political parties developed during the latter stages of his first term, and a two-party system quickly emerged.

A two-party system describes a political environment where two parties dominate elections and typically alternate in power (Janda et al., 2009). Although other parties (third parties) may compete, they rarely win.

Political historians have identified five distinct party systems. The first (1792–1817) featured the Federalists, who were led by Alexander Hamilton and favored strong federal government, and Thomas Jefferson's party (the Democratic-Republicans) who distrusted federal power. After 1800, the Jeffersonians would become the dominant

party since they controlled the presidency and both Houses of Congress. The decline of the Federalists after the War of 1812 culminated in the so-called Era of Good Feelings, since James Monroe was elected president without any opposition in 1820.

By the mid-1830s (1836–1852), a new system emerged. The Democratic Party, led by Andrew Jackson (elected president in 1828), was formed in 1828. Jackson relied on the spoils system, rewarding past supporters and promising jobs to potential allies if they joined his team (Syrett, 1953). He believed that rotation in office would prevent the development of a corrupt civil service.

Those opposed to Jackson called themselves Whigs. Like the Whigs of 1776 who challenged King George III, the new party saw itself as challenging King Andrew. The Whigs supported the American System first proposed by Henry Clay in 1815. This system called for high tariffs to protect and promote American industry, a national bank to provide credit to encourage economic growth, and federal subsidies for internal improvements such as roads and canals that would move agricultural goods to market. The funds for these subsidies would come from tariffs and sales of public lands. Clay believed these policies would allow the United States to become economically independent. Jackson and the Democrats opposed the Whig economic policies, claiming that they favored the wealthy.

For two decades, the parties would compete for power, with the Democrats winning three presidential elections (in 1836, 1844, and 1852) and the Whigs two (in 1840 and 1848).

The election of 1852 would be the last that the Whig Party would contest nationally. The party was ultimately destroyed by the question of whether to allow the expansion of slavery to the territories. Sharply divided on the issue, the antislavery wing blocked the nomination of President Fillmore in 1852 because he had signed the Fugitive Slave Act into law. The Whigs nominated General Winfield Scott, who was defeated by Franklin Pierce. After the election, many of the party's leaders left politics (including Abraham Lincoln) or joined other parties. The party had also been shaken by the deaths in 1852 of Henry Clay and Daniel Webster, two of its longtime leaders. Southern Whigs, known as Cotton Whigs, joined the Democratic Party while Northern Whigs (the Conscience Whigs) moved into antislavery parties.

As the slavery controversy intensified, a number of antislavery political parties appeared, in large part because of the failure of the Democrats and the Whigs to respond to the crisis. Among the parties that appeared during this time were the Liberty Party (1840–1848), the Free Soil Party (1848–1855), the Anti-Nebraska Party (1854), the Opposition Party (1854–1858), and the Constitutional Union Party (1860).

The Republican Party was formed in 1854 by opponents of slavery and supporters of the notion that the federal government should offer free land in the West to settlers.

The new party included abolitionists, who wished to eliminate slavery (who became known as the radical wing of the party), and antiexpansionists (known as the conservative wing), who opposed the expansion of slavery into the territories but were unwilling to outlaw slavery in the states where it existed. Abraham Lincoln, the former Whig congressman, was part of this wing of the party.

By 1860, the Republicans were the principal alternative to the Democratic Party in national politics. Meeting at the so-called Wigwag in Chicago, the Republicans selected on the third ballot a dark horse candidate, Abraham Lincoln, as the party's standard bearer in the general election against a Democratic Party fractured by the slavery issue. The Democrats nominated Senator Stephen Douglas of Illinois and adopted a platform opposing congressional interference with slavery. Southern Democrats, dissatisfied with the platform and Douglas, nominated Vice President John C. Breckinridge and adopted a platform that asserted the following:

That when the settlers in a Territory, having an adequate population, form a State Constitution, the right of sovereignty commences, and being consummated by admission into the Union, they stand on an equal footing with the people of other States, and the State thus organized ought to be admitted into the Federal Union, whether its constitution prohibits or recognizes the institution of slavery. (Porter & Johnson, 1956, p. 31)

Lincoln, who opposed the expansion of slavery into the Western territories but believed that the federal government did not have the authority to ban slavery, was elected, and this critical election (Key, 1955) marks the beginning of the third party era.

During the third party system (1860–1892), the Republican Party would dominate presidential elections, winning six of the eight presidential elections between 1864 and 1892. The only Democrat elected during this period was Grover Cleveland, who was elected in 1884 and 1892. In 1876 (Tilden) and 1888 (Cleveland), the Democratic candidates won the popular vote but lost the electoral vote. Some of the Republican Party's success in presidential elections came from its practice of running Civil War heroes and from the North's domination (until 1877) of Southern politics.

Following the contested presidential election of 1876, President Rutherford B. Hayes removed federal troops from Florida, Louisiana, and South Carolina, the last three states under military occupation. With the end of reconstruction, the Republican Party would virtually disappear from the South until the middle of the 20th century, as the Solid South (as it came to be known) went Democratic.

Although the Republicans dominated the presidency, the Republicans and Democrats won an equal number of congressional elections, each controlling the House of Representatives for nine Congresses between 1860 and 1894. During the period, the Senate, whose members were

still elected by the state legislatures, was dominated by the Republicans, who were the majority in 13 of the 18 Congresses during this time.

It was during this era that the parties' traditional symbols came into use. A political cartoon drawn by Thomas Nast and published in the November 7, 1874, edition of *Harper's Weekly* depicted a Republican elephant and a Democratic donkey; both symbols endure.

The election of 1896 was a critical election because it changed the political environment and resulted in a clear change in the competitive balance between the two parties. The Civil War and reconstruction were no longer salient issues in American politics. The industrial revolution had transformed America, and it would change the competition between the two major political parties.

The country was in a severe economic depression. The Republicans supported big business, the gold standard, protective tariffs, and pensions for Union military veterans. The 1896 Republican platform was also the first to support women's suffrage. The Republican presidential candidate in 1896, William McKinley, won an overwhelming victory over the Democratic candidate, William Jennings Bryan, the so-called free silver advocate who had captivated the Democrats at their convention with his "cross of gold" speech.

During the fourth party era (1896–1930), the Republican Party's domination of the presidency continued. Between 1896 and 1928, only one Democrat, Woodrow Wilson, would win the presidency (in 1912 and 1916). The Republicans also controlled both houses of Congress during 30 of the 36 years between 1896 and 1932. Wilson's first victory, in 1912, was due to a split in the Republican ranks. Theodore Roosevelt, who had become president when McKinley was assassinated in 1901 and was elected in 1904, returned after 4 years to seek the Republican nomination.

Roosevelt, who was a progressive, believed that his hand-picked successor, William Howard Taft, had betrayed the progressive cause by supporting conservative Republican legislation such as the protectionist Payne–Aldrich Tariff Act of 1909. Denied the Republican nomination despite winning most of the primaries, Roosevelt became the nominee of the Progressive Party and came in second, ahead of Taft; the split between the two Republicans allowed Wilson to win. It also gave the Democrats their brief period of control of both houses of Congress during this era.

Although Roosevelt's presidency made his party a progressive force in the early 20th century, by the 1920s, the Republican Party's economic ideology had become promotion of business interests, as exemplified by the party's support of protectionist tariffs. The party also opposed American entry into the League of Nations.

The stock market crash of October 1929 would set the stage for the end of the fourth party era and the end to Republican dominance. President Herbert Hoover, elected

in a landslide in 1928, opposed direct-relief programs to ease the suffering caused by the economic crisis. In the midterm election of 1930, the Democrats narrowed the Republican majorities to one seat in the Senate and two seats in the House of Representatives.

The Great Depression propelled Franklin Delano Roosevelt and his New Deal coalition into power. Roosevelt promised new solutions to unemployment and the economic crisis of the Depression. Roosevelt, who was swept into office in a landslide in 1932, carried huge Democratic majorities with him into Congress to enact the New Deal. With the support of Southern whites, white ethnic voters in big cities, and African Americans, Roosevelt was reelected in 1936, 1940, and 1944.

During the fifth party system (1932–1968), the Democrats were the dominant party, winning 7 of 10 presidential elections and controlling both houses of Congress for all but 4 years.

During this era, there were two wings in the Republican Party. The party's liberals supported most of the New Deal but believed that these programs, as well as other Democratic social programs, could be run more efficiently. They also tended to be internationalists. This wing of the party was geographically centered in the Northeast, and its leading figures included Governor Dewey, his Republican successor Nelson A. Rockefeller, and New York Senator Jacob Javits. The conservative wing of the party opposed the New Deal and, after 1938, joined with conservative Democrats in Congress to block most liberal initiatives until Lyndon Johnson and the Great Society of the 1960s. Senator Robert Taft of Ohio was the longtime leader of this wing, whose base was in the Midwest. During the 1950s, Arizona Senator Barry Goldwater and other conservative Republicans considered bolting from the party to start a new conservative party with conservative Democrats. Goldwater rejected that option and instead sought the Republican presidential nomination in 1964. Although Goldwater lost to Lyndon Johnson in a landslide, he set the stage for the rightward shift of the Republican Party that would culminate in the election of Ronald Reagan in 1980.

During the 1960s, the New Deal coalition began to split over civil rights, the war in Vietnam, and the Great Society programs proposed by President Johnson as part of his war on poverty. In 1968, Richard Nixon was elected president as Hubert Humphrey, the Democrat, would fall victim to Vietnam War opponents and Alabama's Democratic governor, the segregationist George C. Wallace, who carried five states and won 46 electoral votes.

The period since 1968 has been a period of divided government. In the four decades since Richard Nixon's election, one party has controlled the presidency and both Houses of Congress for fewer than 9 years (1977–1981; 1993–1995; January 20, 2001–June 6, 2001; 2005–2007; and 2009–present), with the Republicans winning 7 of 11 presidential elections and the Democrats controlling at least one house of Congress for more than a quarter of a

century (1968–1994), followed by a decade of Republican congressional dominance (1994–2006), with the Democrats retaking control in 2006.

It was also the time when the most significant changes in the selection of presidential candidates in more than 100 years took place. Since the 1830s, party conventions selected the major party's candidates. Following the 1968 Democratic Convention, the McGovern–Fraser Commission proposed changes in the party's rules that made the nomination process more democratic by shifting delegate selection from party leaders (who, through the delegates they selected, controlled the conventions) to voters in primaries and caucuses (the Republicans would make similar reforms in their rules as well). The effect of these changes was to make it possible for party outsiders (notably Jimmy Carter in 1976) to win the nomination and for more candidates to actually pursue their party's nomination in what has become a more drawn out and public process (Asher, 1984).

The Democrat's New Deal coalition began to collapse in the late 1960s as white ethnic and Southern white voters began to drift away from the party. The 1966 midterm election, where the Republicans picked up 4 Senate seats and 47 in the House, was characterized as a white backlash against Johnson's Great Society social welfare policies, the Democratic Party's support of the civil rights movement, and the urban riots. By the late 1970s, the South, which had been a Democratic stronghold since the end of reconstruction, shifted to the Republican column, prompted by President Nixon's so-called Southern strategy (Murphy & Gulliver, 1971).

Ronald Reagan's election in 1980 moved the Republicans sharply to the right. Reagan won a landslide victory over President Jimmy Carter, attracting Reagan Democrats, who were white, Roman Catholic, and blue-collar voters who supported Reagan because of his social conservatism (opposition to abortion rights and gay rights) and his anti-Soviet rhetoric. The 1980 Republican platform, at Reagan's behest, dropped support for the Equal Rights Amendment (which had first been included in the 1940 GOP platform, 4 years before the Democrats would add a similar plank), based on the premise that such an amendment was an intrusion by Congress and the federal courts into a state matter. Reagan was reelected in 1984, winning 49 of the 50 states, and his vice president, George H. W. Bush, prevailed in 1988.

After Democrat Bill Clinton defeated President Bush in 1992, the Republicans won majorities in both houses of Congress in the 1994 midterm elections. Led by House Republican whip Newt Gingrich (who would become speaker), the House Republicans ran on the Contract With America, which promised that, if elected, a Republican majority would bring to the floor bills for a number of reforms, including a balanced budget, welfare reform, and term limits.

For the first time in more than 40 years, the Republicans controlled both houses of Congress (they had

controlled the Senate from 1981 to 1987). For the next 12 years, the Republicans held both chambers (except for the period from June 6, 2001, to January 3, 2003, when the Democrats controlled the Senate after Senator James Jeffords of Vermont left the Republican Party to sit as an Independent).

In 2000, George W. Bush was elected president, despite trailing Vice President Al Gore in the popular vote by 543,816 votes, becoming the first candidate since 1888 to be elected president without winning the popular vote. It was also the first time since 1952 that the Republicans would control the presidency and both houses of Congress simultaneously.

Following the September 11, 2001, terrorist attacks, President Bush gained widespread support as he prosecuted the so-called war on terrorism. The president's popularity helped the Republicans in the 2002 midterm election, allowing them to retake control of the Senate and maintain a majority in the House of Representatives.

Reelected in 2004, Bush's popularity declined because of his failed effort to reform social security, the government's inept response to Hurricane Katrina, and the decline in public support for the Iraq War. The Democrats won back control of Congress in 2006. Barack Obama, the first African American nominated by a major party for president, was elected in 2008 with 365 electoral votes and 53% of the popular vote. Obama carried nine states won by Bush in 2004 and won states such as Virginia and Indiana, which hadn't voted Democratic in more than 40 years.

Third Parties

The Republican and Democratic Parties have dominated American politics for more than 150 years. Since 1852, only Democrats or Republicans have been elected president. Virtually all governors and members of the U.S. Congress are Republicans or Democrats. However, these parties are not the only parties that compete in elections. Third parties have a long history in American politics.

During the 19th century, third parties included the Anti-Masonic Party (1828–1838), the antislavery Liberty Party (1840–1848), the Free Soil Party (1848–1854), the nativist American (“Know Nothing”) Party (1845–1860), the nationalistic Constitutional Union Party (1859–1860), the Greenback Party (1874–1884), the Socialist Party, and the Populist Party (1892–1908). The Prohibition Party (1869–present) is the oldest continuously functioning third party.

In the 20th century, there were a number of third party efforts. The Progressive Party (also known as the Bull Moose Party) was founded in 1912 by supporters of Theodore Roosevelt after the former president (1901–1908) failed to win the Republican nomination. Roosevelt lost, but 17 members of the party were elected to the House of Representatives and one to the U.S. Senate. The 1914

election returned five Progressives to the House. In 1916, the party again nominated Roosevelt. However, he declined the nomination and endorsed Charles Evans Hughes, the Republican candidate. The party then disappeared.

In 1924, another Progressive Party appeared with Wisconsin Senator Robert LaFollette (Republican) as their presidential candidate. Running with Democratic Senator Burton K. Wheeler of Montana, LaFollette polled 16.6% of the popular vote and carried his home state of Wisconsin. After the election, the party disbanded.

In 1948, two factions split from the Democratic Party. The Dixiecrats were Southern Democrats who opposed the Civil Rights plank in the 1948 Democratic Party platform. Their candidate for President, Governor Strom Thurmond of South Carolina, carried four southern states, polling slightly more than 2% of the national vote. The Progressive Party nominated Henry A. Wallace (vice president during Roosevelt's third term) and favored desegregation and maintaining good relations with the Soviet Union. Wallace, who polled around the same percentage of the popular vote as Thurmond, failed to win any states. Twenty years later (in 1968), segregationist Governor George C. Wallace of Alabama would lead the American Independence Party and take enough Democratic votes to help swing the election to Richard Nixon.

In 1980, moderate Republican Congressman John B. Anderson dropped out of the primaries and ran as the candidate of the National Unity Campaign. He received 6.6% of the popular vote, as many of his early supporters shifted to Reagan or Carter.

The Reform Party grew out of H. Ross Perot's 1992 independent presidential campaign. Established in 1995 by Perot as an alternative to the Democratic and Republican parties, the Reform Party emphasized balancing the budget (one of Perot's issues in 1992), term limits for members of Congress, campaign finance reform, and tax reform. In 1996, the party nominated Perot for president and Patrick Choate, an economist and talk-show host, for vice president. The ticket received about 9% of the popular vote (half of the percentage Perot had polled in 1992).

In 1998, Jesse Ventura, a professional wrestler and actor, was elected Governor of Minnesota, becoming the first Reform Party candidate elected to a major office. Ventura and Perot would seek the control the party and the conflict between the two men led Ventura to leave the party in early 2000.

The party has since declined. Patrick J. Buchanan, a conservative commentator, was nominated in 2000 and received less than half a million votes nationally (although many of the 3,400 votes he received in Florida's Palm Beach County were probably intended for Vice President Al Gore, since the county's notorious “butterfly ballot” confused many voters, costing Gore Florida and the election). In 2004, the party endorsed independent presidential candidate Ralph Nader. In 2008, the Reform Party

qualified for the ballot in one state, where Theodore Weill received 470 votes.

Other contemporary third parties include the Libertarian Party (founded in 1971–present) and the Green Party (1991–present).

Third parties have, with the exception of the Republican Party (in 1860), failed to make a breakthrough. The inability of third parties to win elections can be attributed to a number of factors. First, the plurality system favors Democrats and Republicans since third parties rarely attract enough support to win elections. A system of proportional representation, where legislative seats are awarded in proportion to the votes a party receives, would benefit third parties by ensuring that they would have some representation in the government.

Ballot access is another problem for third parties. In most states, only the Democratic and Republican parties are automatically listed on the ballot. Third parties obtain places on the ballot by either their candidate's vote totals exceeding a minimum threshold (for example, in New York State, a party is assured a ballot line for the next four elections if their gubernatorial candidate polls at least 50,000 votes) or if they gather a specified number of petition signatures from registered voters.

Third-party candidates have difficulty raising money and are often ignored by the news media and are excluded from most debates where Democrats and Republicans are routinely invited. Voters often don't even consider voting for third-party candidates because they don't want to waste their vote.

Although third parties have not won a presidential election and rarely win at other levels, they do play a number of important roles in the political process. First, they offer a safety valve for voters who are unhappy with the major parties by giving them a way to express their unhappiness by voting. They also affect policy by introducing issues that the two major parties have avoided; the major parties then often respond out of fear that they will lose votes to the third party. Ross Perot's injection of the deficit and national debt into the 1992 campaign forced the Democrats and Republicans to take up the theme of deficit reduction. Third parties have, at times, helped transform the two-party system. The many third parties that appeared during the 1850s helped pave the way for the third party system; the Populist Party helped reshape the competition between the Democrats and Republicans after the 1896 election, and the American Independent Party, by further distancing Southern whites from the Democratic Party, offered the Republicans an opening during the 1970s that transformed the South from solidly Democratic to solidly Republican.

Future Directions

Political parties and the two major parties are likely to continue to adapt to the changing political environment.

The increasing proportion of the electorate that identifies itself as Independent, coupled with the calls for a postpartisanship by Barack Obama, Arnold Schwarzenegger, Michael Bloomberg, and others (Rauch, 2008), suggest that parties are likely to be weaker in the future.

The rise of the candidate-centered organization has also undermined the parties (Wattenberg, 1991). As candidates employ their own organizations, especially in presidential politics, the national committees (aside from their role in organizing the national conventions) seem less important at that level.

There are other indications that political parties are growing stronger. One such indicator is the incidence of party-line voting in Congress. In the 110th Congress (2007–2009), party-line voting in the U.S. Senate was 84.2%, with 44 of the 102 senators who served in that Congress voting with their party more than 90% of the time (U.S. Congress Votes Database, 2009b). In the House of Representatives, 89.4% of all votes were party line, with 369 of the 448 members who served in the House during this time voting with their party more than 90% of the time (U.S. Congress Votes Database, 2009a). Notwithstanding President Obama's efforts to secure Republican support for his stimulus package, it received no Republican votes in the House of Representatives and 3 (out of a possible 40) in the Senate. In 1970, 58% of House Democrats voted the party line as did 60% of Republicans. Over in the Senate, 56% of the Republicans supported their party's position, and about 52% of the Democrats cast votes on the party line (Janda, 2009). The increase in party-line voting is seen by some as an indicator that parties are growing stronger (Brady, Goldstein, & Kessler, 2002). Others suggest this is more a reflection of the growing ideological polarization of the parties in Congress, with moderate Republicans and conservative Democrats seeing their numbers diminish.

Another indicator of party strength has been the ability of the national committees and each party's congressional campaign committees (the Democratic Senatorial Campaign Committee, Democratic Congressional Campaign Committee, National Republican Congressional Committee, and the National Republican Senatorial Committee) to raise funds, recruit candidates, and offer other forms of assistance in campaigns.

The Bipartisan Campaign Reform Act, with its ban on soft-money contributions to the political parties, was expected to have a substantial impact on the fund-raising activities of political parties. After the legislation was enacted, then-DNC Chair Terry McAuliffe set up a task force to determine how to re-channel the banned soft-money contributions into the campaign process. As a result, a number of so-called 527 committees were set up for the 2004 campaign cycle. In 2004, anti-Bush 527s spent more than \$200 million in their efforts to defeat Bush (Cigler & Loomis, 2006). The Swift Boat Veterans and POWs for Truth attacks on Senator John Kerry's war record certainly damaged his chances of defeating Bush.

The Center for Responsive Politics (2009) analysis of Federal Elections Commission disclosure reports for the 2008 election cycle found that Democratic Party committees raised more than \$961 million and Republican committees collected more than \$920 million. Therefore, although the Bipartisan Campaign Reform Act may have cut off soft-money contributions to the parties, the Democrats and Republicans have adapted their fund-raising efforts to the changes in the law.

As noted earlier, the parties (especially the national committees) have evolved. The national committees have become service organizations, providing an array of campaign services to state and local party organizations and candidates. Some have suggested that this development has revitalized the parties (Aldrich, 1995).

Of course, the parties still retain one important point of relevance to the electorate: They are a cue for the voter. If the voter knows nothing else about the candidate (which is often the case with candidates in down-ballot elections), party affiliation offers some guidance as to where the candidate stands on the issues.

As for the future of the two-party system, the 2008 election may turn out to have been a critical election. Barack Obama's historic victory, with the Democrats expanding their majorities in both the House and Senate, creates the opportunity for a new period of political domination by the Democratic Party (Sabato, 2009).

However, history has demonstrated that the two major parties are remarkably resilient, and the significance of the 2008 election will, more than likely, be based on the success that the Democratic president and his colleagues in Congress have in reversing the nation's economic decline and keeping the nation secure.

Conclusion

Although not established in the Constitution, political parties have become essential to the American political process. Their roles in elections and governing have evolved over time, and some have questioned the continuing viability of our parties. However, others argue that the move from party-centered to candidate-centered elections has revitalized the parties and that they continue to remain an important part of our political landscape.

The parties play important roles in elections and governing and stabilizing our political system. The two-party system has existed almost as long as the republic, with the same two parties (the Democrats and Republicans) dominating the political landscape for more than 150 years. Third parties, while rarely an electoral threat, provide those dissatisfied with the two parties opportunity to use the electoral process to raise issues and question the system.

Elections have changed, and the parties, rather than withering away as some have predicted (or hoped), have

continued to evolve and remain critical to the functioning of American democracy.

References and Further Readings

- Aldrich, J. H. (1995). *Why parties? The origin and transformation of political parties in America*. Chicago: University of Chicago Press.
- Allen, O. E. (1993). *The tiger: The rise and fall of Tammany Hall*. Reading, MA: Addison Wesley.
- American Political Science Association Committee on Political Parties. (1950, September). Toward a more responsible two party system. *American Political Science Review*, 44(3).
- Asher, H. B. (1984). *Presidential elections and American politics* (3rd ed.). Homewood, IL: Dorsey.
- Berman, A. (2008, November). The redemption of Howard Dean. *Guardian*, 11. Available at <http://www.guardian.co.uk/commentisfree/cifamerica/2008/nov/11/barack-obama-howard-dean-democrats>
- Brady, D., Goldstein, J., & Kessler, D. (2002). Does party matter? An historical test using Senate tariff votes in three institutional settings. *Journal of Law, Economics, and Organization*, 18(1): 140-154.
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. New York: Wiley.
- Campbell, A., Gurin, G., & Miller, W. E. (1954). *The voter decides*. Evanston, IL: Row, Peterson.
- Center for Responsive Politics. (2009). *Political parties: Overview of 2008 election cycle*. Available at <http://www.opensecrets.org/parties/index.php>
- Cigler, A. J., & Loomis, B. A. (2006, August-September). *James Madison, E. E. Schattschneider, and Tom DeLay walk into a bar: organized interests, political parties and the politics of representation*. Paper presented at the annual meeting of the American Political Science Association, Philadelphia, PA.
- Conley, B. (2008, January). *Route to '66: Ray Bliss and the 1966 mid term Republican resurgence*. Paper presented at the annual meeting of the Southern Political Science Association, New Orleans, LA.
- Democratic Party of the United States. (2007). *The charter and the bylaws of the Democratic Party of the United States, as amended by the Democratic National Committee, February 3, 2007*. Washington, DC: Democratic National Committee.
- Downs, A. 1957. *An economic theory of democracy*. New York: Harper & Row.
- Duverger, M. (1964). *Political parties: Their organization and activity in the modern state* (2nd ed., B. & R. North, Trans.). London: Methuen.
- Huckshorn, R. (1984). *Political parties in America*. Monterey, CA: Brooks/Cole.
- Janda, K., Berry, J. M., & Goldman, J. (2009). *The challenge of democracy: American government in a global world* (10th ed.). Boston: Wadsworth Cengage Learning.
- Key, V. O. (1955). A theory of critical elections. *Journal of Politics*, 17, 3-18.
- Morehouse, S. M., & Jewell, M. E. (2003). *State politics, parties and policy* (2nd ed.). Lanham, MD: Rowman & Littlefield.
- Murphy, R., & Gulliver, H. (1971). *The Southern strategy*. New York: Scribner.

- Pew Research Center for People and the Press. (2009). *Independents take center stage in the Obama era: Trends in political attitudes and core values: 1987-2009*. Washington, DC: Pew Research Center.
- Porter, K. H., & Johnson, D. B. (1956). *National party platforms: 1840-1956*. Urbana: University of Illinois Press.
- Rauch, J. (2008, July-August). Post partisanship. *The Atlantic Monthly*, 301, 6.
- Rosenblum, N. L. (2008). *On the side of the angels: An appreciation of parties and partisanship*. Princeton, NJ: Princeton University Press.
- Rosenthal, C. S. (1995). New party or campaign bank account? Explaining the rise of state legislative campaign committees. *Legislative Studies Quarterly*, 20, 249-268.
- Sabato, L. J. (2009). *The year of Obama: How Barack Obama won the White House*. New York: Pearson Longman.
- Schattschneider, E. E. (1942). *Party government*. New York: Holt.
- Schneider, W. (1992). The suburban century begins. *Atlantic Monthly*, 270(1), 33-44.
- Shea, D. M. (1996). *Campaign craft: The strategies, tactics, and art of political campaign management*. New York: Praeger.
- Sifry, M. L. (2003). *Spoiling for a fight: Third party politics in America*. New York: Taylor & Francis.
- Stonecash, J. (2005). *Political parties matter: Realignment and the return of partisan voting*. Boulder, CO: Lynne Rienner.
- Syrett, H. C. (1953). *Andrew Jackson: His contribution to the American tradition*. Indianapolis, IN: Bobbs Merrill.
- The U.S. Congress votes database. (2009a). 110th Congress, House. *The Washington Post*. Available at <http://projects.washingtonpost.com/congress/110/house/party-voters>
- The U.S. Congress votes database. (2009b). 110th Congress, Senate. *The Washington Post*. Available at <http://projects.washingtonpost.com/congress/110/senate/party-voters>
- Washington, G. (1940). *The writings of George Washington from the original manuscript sources 1745-1799* (Vol. 35, J. C. Fitzpatrick, Ed.). Washington, DC: Government Printing Office.
- Wattenberg, M. P. (1991). *The rise of candidate centered politics*. Cambridge, MA: Harvard University Press.
- West's Annotated Code of Virginia* (Vol. 19). (2003). Eagan, MN: West Group.
- White, J. (2009). *Political parties*. Available at <http://www.apsanet.org/content/5221.cfm>

90

STATE AND LOCAL GOVERNMENT

JUNE SAGER SPEAKMAN

Roger Williams University

The study of state and local government is essentially the study of all that is not national government in the United States—the 50 states and the more than 88,000 other subnational units of government from counties to small towns, fire districts, school districts, and water districts. Typically, the study of cities—communities of larger than 50,000 people—is considered a separate realm of inquiry (see Chapter 81, titled “Urban Politics”). The distinction among these layers of government has been confounded in recent years with the emergence of the metropolis—such as the so-called BosWash region that encompasses the Northeast corridor from Boston to the nation’s capital and includes all the cities, suburbs, and rural areas in between.

This chapter examines where state and local government sit in the American national political system and then focuses on the internal workings of each level, looking at executives, legislatures, judicial systems, political participation, and policy making.

States in the Federal System

State governments, and the municipalities within them, preceded the creation of the national government. A drive through the northeastern United States will reveal cities and towns founded in the 1600s and 1700s, long before the 1787 writing of the founding document of the nation in

which they sit. Thirteen states were viable political entities at the time of the American Revolution. The states had their own constitutions, forms of government, political processes, political cultures, and political identities. Virginians, New Yorkers, and Pennsylvanians existed long before Americans.

The preexistence of states had a profound impact on the design of the Constitution and continues to shape the operation of American politics today. The president is elected by electoral votes, which are allocated to states. This unique electoral institution determines presidential campaign strategy and, some argue, has an impact on presidential policy making as the president responds to the states that are vote rich. Both chambers of Congress are organized geographically, with Senators originally chosen by state legislators (until the Seventeenth Amendment in 1913, which ushered in direct election) and House members elected from districts within states. The shape of these districts is determined by state legislatures, creating one of the most politically charged elements of American politics, as state legislatures use their line-drawing powers to the benefit of the party that controls that legislature. The Constitution itself can be amended only with the approval of three fifths of the state legislatures, making constitutional change by this method rare indeed.

Clearly, state governments matter a great deal in American politics. The impact of structural elements outlined is enhanced by the constitutional allocation of power

between the state and federal governments. Articles I, II, and III of the Constitution allocate certain powers to the national level of government. Many powers and many areas of authority are not mentioned in the provisions and have been, over the course of 220 years of constitutional history, contested terrain between state power and national power. The Tenth Amendment to the Constitution, often called the states-rights amendment, asserts that all powers not granted to the national government or denied to the states “are reserved to the States respectively, or to the people” (U.S. Constitution, 1787, Amendment Ten). This reserved-powers clause has been used by state governments with considerable success to fend off national incursion into state sovereignty. In 1997, for example, in *Printz v. U.S.*, the Supreme Court ruled unconstitutional a provision of the Brady Bill gun control bill that required local police to conduct background checks for handgun purchases. The Court saw this provision as a violation of the federal division of powers. Also in 1997, a federal district court affirmed the right of the voters of California when they passed Proposition 209 banning affirmative action by public institutions in the state. The proposition is in direct contraction to federal policy but was allowed to stand by the Court. Not all attempts to assert state sovereignty succeed. In 1985, in *Garcia v. San Antonio MTA*, the Court added to a long line of decisions, beginning with the New Deal, that permitted federal regulation of the conditions of labor for state and local employees, including minimum wage, maximum hours, and the right to unionize.

In the early 21st century, the boundaries between state and federal authority remain unclear. In the area of medical marijuana, for example, at least 14 states have laws that permit the cultivation, distribution, and use of marijuana for medicinal purposes, and several have decriminalized marijuana altogether. However, marijuana remains illegal at the federal level, and there have been numerous incidents, notably in California, where medical marijuana dispensaries, operating in a manner consistent with state law, have been shut down by federal authorities. In the 2005 *Gonzales v. Raich*, the Supreme Court asserted the federal government’s power to prohibit the use of marijuana even for medicinal purposes. This particular federal–state dispute remains active as more states indicate the likelihood of relaxing marijuana laws at the state level.

States remain vibrant, vital sovereign actors in American politics. Constitutional provisions, centuries-old practice, and Supreme Court decision making will ensure that state–federal tension remains a significant dynamic in the system.

Variety and Inequality

One of the frequently cited benefits of an American-style federal system is the variety offered by the existence of 50 sovereign states. From the physical differences between

the Alaska tundra and the Hawaiian tropics, to the cultural differences between Utah and Florida, to the economic differences between Mississippi (median household income of \$37,700) and Connecticut (median household income of \$70,500) and the political differences between Rhode Island (63% voted for Obama) and Wyoming (33% voted for Obama), diversity is an undeniable fact of American life. This diversity provides a number of benefits. It provides an array of opportunities and options for American citizens in terms of economic opportunity and lifestyle choices. Political scientist Daniel Elazar (1984) observed three distinct political cultures in the United States. His successor in the field, Joel Lieske, identified 10 regional sub-cultures (Bowman & Kearney, 2008). These allow for the prediction of quality of life, business climate, expansiveness of government programming, and other features.

Anyone traveling around the United States can observe differences in speed limits, motorcycle helmet laws, hours of operation of bars, and severity of punishment for crimes (three-strikes laws and the presence of a death penalty). These obvious differences among states are the surface manifestations of the considerable leeway that states have to determine policy agendas.

Many observers note that some of the most significant policy innovations in contemporary American history have begun as experiments at the state level. Welfare-to-work programs, universal health insurance, charter schools, and family leave policies are in this category.

As noted, however, the presence of state autonomy with regard to public policy can produce inequities among Americans that derive from the state in which they live. Federal laws such as the No Child Left Behind Act attempt to mandate a national standard for education. Federal programs such as Food Stamps and Medicaid ensure a base level of benefits to the poor regardless of geography. Even with these programs, there is considerable variation in the funding levels and quality of services from state to state.

State Political Structures

State Constitutions

Article IV of the U.S. Constitution lays out several rules that must be followed by all states if they are to be members of the union: Each state shall give full faith and credit to the records and documents of every other state, citizens of every state shall have all the privileges and immunities granted to citizens of each state, and every state shall have a republican form of government. Article I, Section 10 prohibits states from doing certain things: coining money, entering into treaties, passing ex post facto laws, among others. Beyond these relatively few requirements, states are free to design their own forms of government.

Although there is some variety from one state to the next in constitutional structures and processes, most states

mimic the national level of government, with three branches, checks and balances, regularly scheduled elections, substate units of government with some degree of autonomy, and some set of civil rights and liberties guaranteed to the citizens.

Beyond those basic structural similarities lie significant differences among state constitutions. Massachusetts has the oldest constitution, adopted in 1780. Georgia, on the other hand, adopted a new constitution in 1983 (Saffell & Basehart, 2001). Some constitutions are very long, like Alabama's with over 340,000 words, some are rather short, like New Hampshire's with 9,200 words, although none as short as the U.S. Constitution with 8,700 words (Bowman & Kearney, 2008). Most state constitutions are relatively easy to amend, including by popular initiative. In 17 states, the constitution can be amended by majority vote of the legislature. In 18 states, the constitution can be amended by majority vote of the electorate. California's Proposition 209, noted previously, is one example of this. Thirty-one states have constitutional amendments restricting the definition of marriage to a union between one man and one woman. The result is a proliferation of provisions attached to many state constitutions. In 2002, a provision was added to the Florida constitution that prohibits "cruel and unusual confinement of pigs during pregnancy" (Bowman & Kearney, 2008, p. 59). Seven constitutions have bills of rights that make mention of dueling (Saffell & Basehart, 2001).

One of the most controversial recent uses of the state constitution amendment process has been in the area of gay marriage. In the United States, demands for equal treatment for all citizens regardless of race, gender, religion, age, and disability have been made and resolved, for the most part, at both the state and national levels. Such demands for equality in the matter of sexual orientation are still contested terrain in U.S. politics. Family law is typically the province of state governments, and so the battle about the legality of same-sex marriages is being fought in that arena. Since state courts, state attorneys general, and local town clerks have frequently found no basis in law for denying a marriage license to same-sex couples, many states have responded by defining marriage as between one man and one woman either by statute (in 12 states) or by an amendment to the state constitution (in 29 states). Virtually all of these amendments have been adopted since 2000. Most notably, in California, in November 2008, voters approved Proposition 8, which amended the state constitution to define marriage as between a man and a woman. This vote came only months after a California court had ruled same-sex marriage as constitutional in that state. So the voters changed the constitution, redefining the rights of gay Californians (Archibold & Goodnough, 2008). The ease with which many state constitutions can be amended by both legislatures and voters have resulted in some constitutions with hundreds of amendments: South Carolina with 484, California with 848, and Alabama

with 1,028. The result is constitutions that are weighed down with anachronistic policy mandates; confusing, overlapping, and conflicting prohibitions; and special protections for groups that have been savvy enough to take advantage of the amendment process.

Starting in the 1960s and continuing to this day, most states have been undertaking a process of constitutional reform to address some of the flaws of these wordy, policy-laden documents. Bowman and Kearney (2008) reveal that between 1960 and 1980, every state altered its constitution to some degree, and 10 states replaced their constitutions completely. The process of reform, which is connected to similar trends in other areas of state government to be discussed in subsequent sections, continues.

Governors

The governor is the chief executive of the state, charged with the day-to-day operation of the state. Yet the aversion to executive power that suffuses the national government is present at the state level as well. In early state constitutions, governors were often limited to single terms, or even one-year terms, and had no veto power, no budgeting power, and no appointment power. In some states, the governors themselves were appointed by the legislature. As with the national government, as the states grew more complex, the need for effective executive power became clear. And throughout the 19th century, state legislatures revealed themselves to be prone to corruption, and their oversight of the budget and administration of state government became problematic. The reform movement of the Progressive Era led to some changes in the allocation of power at the state level as governors' terms were lengthened and many were given the veto power. The first half of the 20th century was a period of little change in state government as the federal government grew rapidly and, with the New Deal, gathered much tax, spending, and programmatic power to Washington, D.C.

Throughout this period, right through the mid-1960s, the malapportionment of state legislatures rendered most governors powerless in the face of their legislatures. Virtually all state legislative lines were relics of earlier times, drawn before populations of immigrants and farm workers swelled the size of American cities. So by the early 1960s, state legislatures were dominated by representatives from tiny rural districts. Governors, on the other hand, were elected statewide in response to the needs of the population centers. But without legislative approval, governors were unable to push their policy agendas. In Tennessee, one Charles Baker of Memphis pointed out to the Supreme Court that his district, with its one representative, had 10 times as many people in it as neighboring rural districts. This underrepresentation of urban districts in both state legislatures and the U.S. House of Representatives was ruled a violation of the equal-protection clause of the U.S. Constitution.

The 1962 *Baker v. Carr* decision was crucial in the reallocation of power within state legislatures from the rural areas to the population centers and removal of legislative barriers to the governors' ability to make policy (Weber & Brace, 1999). At the same time, changes at the federal level were pushing more decision-making authority and some money down to the state level.

Starting in the 1970s with President Nixon's New Federalism and continuing on through Presidents Reagan, Bush, and Clinton, the so-called devolution revolution redesigned federal programs by loosening up guidelines on how federal grant money could be spent at the state level. Categorical grants in which Congress determined the shape of programs were redesigned as block grants in which governors could choose how to spend the federal dollars within loose national guidelines. Governors became important policy actors with dramatically enhanced responsibility for program design and implementation.

Again, the formal powers of a governor vary from state to state, but it is the case that the overall status and importance of the governorship has been dramatically enhanced since the founding of the American republic. The significance of the office can be seen by the number of governors who have gone on to become presidents in recent years—Carter, Reagan, Clinton, and George W. Bush—and the number of governors (four) who have been tapped to serve in the Obama administration—Vilsack of Iowa at the Department of Agriculture, Locke of Washington at the Department of Commerce, Sebelius of Kansas at the Department of Health and Human Services, and Napolitano of Arizona at the Department of Homeland Security.

Because governors are both politicians and administrators, with a wide range of policy areas as their responsibility, and are on duty in their states around the clock, they are seen by many experts as more likely to be better prepared for the presidency than are members of Congress. Most scholars agree that today's governors are the following:

better educated, more experienced in state government and more competent than their predecessors. Never before has the strength and policy influence of the governors been surpassed. . . . The governors have displayed greater capability and vigor than ever before. (Bowman & Kearney, 2008, p. 195)

State Legislatures

As with governors, it is difficult to generalize about state legislatures except to say that most are part-time, all but one are bicameral, and all serve the functions of representation, legislation, and oversight of the executive branch (Bowman & Kearney, 2008). The National Council of State Legislatures has divided states into categories based on length of session. Red states have legislatures that are in session for at least 80% of the work year. Those

serving in these bodies earn a salary that allows them to be full-time legislators. It is clear that larger states, with complex policy needs, fall into this category. In 2006, a California legislator earned \$110,000, while a legislator serving in New Hampshire earned \$200.

In terms of processes, state legislatures work much the same way that the U.S. Congress does. Members are elected from districts and must attend to the needs of their constituents. Yet they must also participate in the making of state policy by serving on committees, writing and voting on legislation, and interacting with the governor and the state bureaucracy. As with the governor, as demands put on states have increased since the 1960s, many state legislatures have become increasingly professionalized by adding staff, lengthening sessions, and raising salaries. In some states, this had led to the same kind of incumbency effects that exist at the national level as some legislators hold on to their seats for decades. In reaction to this, throughout the 1990s, voters in 15 states adopted initiatives that limited state legislatures' terms in some ways. Term limits in the state legislatures were expected to bring in new blood, eliminate entrenched interests, and make the body more representative of minority groups in the population that had been locked out by incumbents. According to several studies, however, term limits in state legislatures have produced a slower-working institution, a stronger governor, more influential interest groups, and term-limited legislators who are less, not more, responsive to constituents. Further, there has been no measurable increase in the representation of underrepresented groups in state houses.

What is clear is that important policy decisions are made in state capitals. Political scientists must attend to the analytical and decision-making capacity of both governors and state legislatures. The challenges facing state governments are enormous since many face large deficits, job losses, crumbling infrastructure, flawed education systems, and other major policy problems. Is the capacity to address these problems present in the states? The answer to that question will be revealed in the early decades of the 21st century.

State Judicial Systems

State courts process 100 million cases a year, which is 99% of all the cases heard in American courts (cited in Weber & Brace, 1999). Clearly it is state courts that are the heart of the U.S. judicial system. As with governors and state legislatures, state courts have changed significantly over the past 40 years, professionalizing and reforming to enhance their efficiency and legitimacy. State courts deal with a wide range of issues that affect the daily lives of all Americans: family law, traffic law, property disputes, debt collection, and criminal law. State courts are also usually the first to look at cases that can have monumental constitutional import, from whether to allow prayer in public

schools to the appropriateness of race-based admissions in state law schools to whether local governments can seize private property for economic development purposes.

As with legislatures and executives, there is considerable variability among the structures and processes of state judicial systems. Some states have multiple types of courts, some have elected judges, some have mandatory sentencing, others have three-strikes laws, and still others have considerable judicial discretion regarding the disposition of cases.

One of the major structural features that distinguish one state court system from the next is the method for selecting judges. For many years, starting with Mississippi in 1832 and continuing throughout the 19th century as new states entered the union, the judicial selection process of choice was election by the state legislature or by election of the voters. Both of these methods came under criticism during the Progressive Era as reformers voiced concern about the ability of elected judges in particular to have the qualifications, objectivity, and accountability to make sound judicial findings. In 1937, the American Bar Association introduced and endorsed the Missouri Plan, which involved a judicial nominating commission recommending judicial nominees to the governor. The governor would then appoint those recommended (should he choose to); the appointees would take office and then be ratified by the voters in a so-called retention vote at the next regularly scheduled election. Twenty-three states use some form of this method (not always including the retention vote), 22 continue to use elections, and 5 use pure gubernatorial appointment. As with the trends in constitutional, executive, and legislative reform, state judiciaries have followed a pattern of increasing professionalism and modernization (Bowman & Kearney, 2008).

Bureaucracy

One of the characteristics of state government that does not mimic federal government is the so-called long ballot, which provides for the election of numerous statewide officials, who, at the federal level, would be appointed by the executive. This feature is a reflection of the early suspicion of executive power, and despite numerous attempts at reform, it has remained an element of most state governments. So the heads of major state departments—the attorney general, the state treasurer, and the commissioner of education—are often elected by the voters and immune to gubernatorial control. Although this allows for greater responsiveness to the public, it does create opportunities for gridlock or redundancy in policy making.

State bureaucracies and their counterparts at the local level are the agencies that deliver virtually all public goods and services, from education to transportation to policing, parks, waste management, and water supplies. The precise form of this service-delivery mechanism varies significantly from state to state, with some states

relying on strong county government structures to deliver most services and others relying on local governments at the point where the rubber meets the road. Some states are complex systems of elected or appointed boards and commissions to advise on or make policies, while others are not. In 2008, state governments employ about 3.8 million (down from 4.6 million people in 1992). Of these, 1.2 million work in higher education, 477,000 in corrections, and 377,000 in corrections (U.S. Census, 2009).

In the last half of the 20th century, the functions of government have become more complex, and the devolution revolution has shifted responsibility for delivering services from the federal government to the states. Along with these developments has come increased attention to the need for efficiency and transparency in the delivery of state services. As with governors and legislatures, state bureaucracies have, over the course of the past several decades, undergone a wave of reform and professionalization. One of the elements of reform has been consolidation. In 1993, for example, South Carolina reduced the number of state agencies from 79 to 17 and eliminated many boards and commissions (Saffell & Basehart, 2001). Many states have adopted some form of performance budgeting, in which agencies' budgets are tied to the quality of services delivered. Some states have looked to privatization of some services as a means of improving service delivery and saving money. And virtually all state governments have moved in significant ways into so-called e-government, eliminating much of the paperwork and waiting lines that had been the source of public frustration at such agencies as motor vehicle registries and public assistance agencies.

Each year, *Governing* magazine, in collaboration with the Pew Center on the States, grades the 50 states on their performance in four areas of public management: budgeting, personnel management, infrastructure, and management of information. Detailed reports on the management strengths and weaknesses of each state reveal dramatic differences in administrative efficiency, modernization, citizen satisfaction, and quality of services delivered. The three states receiving the highest overall grade (A–) in 2008 are Utah, Virginia, and Washington. The two states at the bottom are Rhode Island (C–) and New Hampshire (D+). Students of state politics and public management can find a wealth of information and much fodder for future research in these reports (Pew Center on the States, 2008).

Local Governments

County Government

There are 88,000 units of local government in the United States. Of these, 3,033 are county governments. The size and functions of counties vary dramatically from state to state. Counties employ 2 million employees nationwide and range in size from 6 square miles (Arlington

County, Virginia) to 87,000 square miles (North Slope Borough, Alaska) and population from 67 residents in Loving County, Texas, to 9.5 million residents in Los Angeles County, California. In Rhode Island and Connecticut, counties exist as lines on a map but nothing else. Virginia relies on counties to deliver most educational services in the state, while New Hampshire uses county governments to deliver most public welfare services. Counties are particularly important units of government in rural areas where residents may live in areas that have not been incorporated into a city or town. For these citizens, county government is their primary contact point for all public goods and services. In keeping with the movement to reform government at the state and local levels, many states are examining their county governments with an eye to consolidation and streamlining. In California, for example, the legislature considered a proposal to consolidate the state's 58 counties into seven regional supergovernments. The bill did not pass, but it is an indication of current thinking about the place of county governments in the United States (Bowman & Kearney, 2008).

Municipal Government

According to the U.S. Census (2009), in 2008, there were 19,492 municipal governments in the United States, 16,519 town governments, 13,051 school districts, and 37,381 special district governments. Combined employment in this sector was 14.2 million people. Again, the primary characteristic of municipal government is diversity. The United States Constitution was silent on the question of local governments, leaving to the states to determine how citizens would be represented and services delivered in their states. Virtually all states have some version of home rule, in which cities and towns are granted some degree of autonomy over their own affairs. There is, however, a long tradition of state interference with local affairs that has been enshrined in Dillon's rule, expounded by Iowa judge John Dillon in 1868, which states that local governments have only those powers explicitly granted to them by the state government. Although this is contrary to the Jeffersonian principle that the most desirable form of government is that which is closest to the people, it is the guiding tenet of state-local relations to this day (Bowman & Kearney, 2008).

Municipal governments can take several forms. Larger cities in the Northeast and Midwest tend to organize according to the strong mayor-council form of government, which mimics the state and federal models. Mayors are elected citywide and are responsible for the day-to-day operation of the city. City councils are typically part-time and responsible for enacting legislation, including the budget, connecting with the residents (since the council is usually elected from wards, or districts, in the city), and exercising oversight of mayoral activities. Smaller communities often use the council-manager form of government in which the town council is elected, either at large or from

districts, and is very part-time with limited powers. The council appoints the town manager, or administrator, who is a professionally trained (in the best case) public manager. He or she has authority over all town departments and is responsible for the daily operation of the town.

Most local governments also have a number of special-purpose districts or authorities. These bodies are usually governed by an elected board that oversees an appointed professional executive. Education, sewers, bridges and tunnels, fire service, ports, airports, libraries, and many other service areas can be operated by these commissions. The commissions are often formed to extract the policy area from the control of the town government for reasons of efficiency or to deliver a service that involves multiple jurisdictions. These districts have been criticized in recent years for lack of transparency and accountability and redundancy of services.

The proper way to organize municipal government and deliver local services is not a settled matter. Students of political science and public administration can find many intriguing questions in this area of inquiry. The local level is the place where citizens have the most frequent and most direct contact with their government. It is also the level where political scientists have the best access for research and where those who wish to combine theory and practice can most easily do so. Much can be learned about the dynamics of elections, the operations of bureaucracy, the nature of executive power, best and worst practices in public policy from observation, and analysis of those units of government close to where we live.

Because of the extraordinary diversity in political structures, processes, and policies from state to state and municipality to municipality, comparative analysis is easy to do and bears significant fruit for the student of government.

Conclusion

Although there is only one national government in the United States and that receives the lion's share of attention from political scientists and the media, there are 50 state and approximately 88,000 local units of government. The existence of these state and local governments is enshrined in the U.S. Constitution and in American political culture. Their existence provides an impediment to effective national policy making and an opportunity for diversity and experimentation at the state and local levels. Their existence provides a multitude of arenas for public participation in politics and close-up observation of the processes and politics of governance. Students of political science can learn much about executive power from studying governors, much about legislative power from studying city councils, and much about grassroots activism from studying campaigns for mayor or the multitude of local policy initiatives that appear on ballots at election time. Recent commitments to transparency, the recent

trends in e-government, and the physical proximity of the units of government provide ready access to the researcher. These so-called laboratories of democracy are also laboratories for the discipline.

References and Further Readings

- Archibold, R. C., & Goodnough, A. (2008, November 5). California voters ban gay marriage. *The New York Times*. Retrieved February 15, 2010, from <http://www.nytimes.com/2008/11/06/us/politics/06ballot.html?scp=3&sq=california%20gay%20marriage%20ban&st=cse>
- Bowman, A., & Kearney, R. (2008). *State and local government* (7th ed.). Boston: Houghton Mifflin.
- Council of State Governments: <http://www.csg.org>
- Elazar, D. (1984). *American federalism: A view from the states*. New York: HarperCollins.
- Engel, M. (1999). *State and local government*. New York: Peter Lang.
- Governing: <http://www.governing.com>
- Human Rights Campaign. (2010, January 13). *Statewide marriage prohibitions*. Retrieved February 15, 2010, from http://www.hrc.org/documents/marriage_prohibitions_2009.pdf
- National Association of Counties. (2009). *An overview of county government*. Retrieved February 16, 2010, from [http://www.naco.org/Content/NavigationMenu/About Counties/County Government/A Brief Overview of County Government .htm](http://www.naco.org/Content/NavigationMenu/About%20Counties/County%20Government/A%20Brief%20Overview%20of%20County%20Government.htm)
- National Council of State Legislatures. (2009). *Full and part time legislators*. Retrieved February 15, 2010, from <http://www.ncsl.org/?tabid=16701>
- National Governors Association: <http://www.nga.org>
- Olshfski, D., & Cunningham, R. (2008). *Agendas and decisions*. Albany: SUNY Press.
- Pew Center on the States. (2008). *Grading the states*. Retrieved February 16, 2010, from http://www.pewcenteronthestates.org/gpp_report_card.aspx
- Saffell, D. C., & Basehart, H. (2001). *State and local government*. Boston: McGraw Hill.
- Stateline.org: <http://www.stateline.org>
- U.S. Census. (2009). *State government employment data: March 2008*. Retrieved February 16, 2010, from <http://www2.census.gov/govs/apcs/08stus.txt>
- U.S. Conference of Mayors: <http://www.usmayors.org>
- U.S. Constitution. (1787). Available at <http://www.usconstitution.net/const.html>
- Weber, R., & Brace, P. (Eds.). (1999). *American state and local politics*. New York: Chatham House.

91

PUBLIC POLICY AND ADMINISTRATION

CANDACE C. YOUNG

Truman State University

Policy studies emerged as an important focus in political science in the 1970s. In 1969, David Easton (1969), president of the American Political Science Association, was frustrated with the trend in political science research to study narrow questions that lent themselves to the quantitative methods expected by the behavioral movement. Thus, he called for a post-behavioral revolution where political scientists would study the most important political problems of the day even when quantitative methodologies could not be employed. Easton's call served as a catalyst for policy research that sought to explain and predict policy patterns as well as to evaluate the relative impact of various types of policy solutions.

This chapter discusses a variety of approaches to the study of public policy and shows how the public administration and public policy subfields are closely related and at times intertwined. At the time of Easton's call for relevance, the public administration subfield had declined as a prominent subfield in the discipline. The behavioral movement had prevailed in expectations for quantitative research, and public administration had not moved toward a grand theory or wed itself to quantitative methods. However, it had gravitated toward more policy-relevant models and concepts that were important foundations for the emerging field of public policy.

One of the policy subfield's great advantages for those interested in government and politics is its interdisciplinary

and holistic focus. Research is broad ranging and borrows heavily from the work of neighboring subfields and other academic disciplines. This can be found in its use of ideas such as systems theory, plural and elite models, subsystems concepts, and decision-making research.

Over the last four decades, a number of strands of policy research have developed as the subfield has matured. This chapter uses the policy stages framework to organize most of the literature discussed. However, a few of the newer policy frameworks follow this discussion in the alternative models and future directions sections of the chapter.

Public Administration as a Foundation for the Study of Public Policy

Policy and administration studies have many areas of overlapping interest. Michael Nelson (1977) suggested that the popularity of policy studies was temporal and it was more a modified version of public administration than a new subfield. A close study shows that the two fields do have substantial overlap, and a review of public administration literature is important to one's understanding of public policy.

American scholarly attention to public administration dates back to the late 1800s when Woodrow Wilson (1887) wrote his classic essay calling for the development of a career public service. An increase in the professionalism of

government administration was necessary to meet the increasing quantity and complexity of government activities. He encouraged comparative study of administration and argued that since administration is distinct from politics, the United States could examine administrative practices of European monarchies without fear of undermining its democratic form of government.

Wilson's work was written around the same time that the path-breaking German sociologist Max Weber conducted his seminal studies on bureaucracy. Weber (1946) and Wilson (1887) each posited principles of efficiency, centralized authority, hierarchical structures, educated workers, and application of expertise to administration. Weber saw the development of bureaucracy as a natural corollary to modern government and asserted that its rule-driven decisions supported the rule of law and egalitarian values of democracy. The classical model of administration was further developed by scholars who participated in the scientific management and principles schools of administration. Scholars such as Luther Gulick, Frederick Taylor, and Leonard White reinforced the view of bureaucracy as a rational, efficient, hierarchical machine. This paradigm contributed to the sense that administration and policy were conducted in separate spheres and that organizations were controlled by the administrator at the top of the organization's hierarchy. A careful reading of the early scholars, especially Wilson and Weber, shows that they realized the line between policy and politics was not as distinct as later scholars' attributions. Wilson (1887) discussed the need for public opinion to be a guide for administrators but also stated that administrators should have some discretionary authority. Weber (1946) cautioned that bureaucrats would use their wealth of information and knowledge to their advantage, observed that bureaucrats were likely to categorize specialized information as official secrets, and warned that an authoritative monarch would be powerless opposite an administrative expert.

Open-Systems Administration

By the mid-1900s, many administration scholars challenged the classical model and its primary attention to structure, formal rules, and hierarchy within a single organization. Instead, open-systems scholars discussed the influence of other systems on the political system and how changes in the environment required organizations to adapt. Philip Selznick (1949) in his study, *TVA and the Grass Roots*, revealed how significantly local grassroots organizations and interests can affect an agency implementing public policy. He showed how organizations have to consider threats from external organizations and interests. One strategy to lessen or neutralize the threat was co-optation. Organizations incorporated dissident parties either formally or informally into their decision-making structures. These

representatives provided increased legitimacy by expanding the perspectives that made up the decision-making body. Ideally, the representative also communicated information favorably back to the external group. When necessary, agency officials changed policy requirements to reduce external hostility to their programs. Since these policy changes occurred without participation of elected officials, a more positive view of co-optation suggested that it increased the level of democratic participation at the local level. Selznick's contributions to an understanding of the important role of external influences, implementation, and inter-governmental complexity have been significant contributions to the study of public policy.

In addition, the open-systems model encouraged thinking about organizations as organisms rather than human machines. Thus, to understand organizations, scholars need to study both formal and informal elements rather than rely on the overwhelming emphasis that the classical model places on formal structures. Chester Barnard (1938) posited that executives and scholars must seek to understand an organization's people, customs, myths, and values as much as the organization's structure and rules.

The open-systems model of administration continues to contribute to policy scholarship. It helps to show that policies are not self-implementing and that the administrative variable has an independent impact on the effectiveness of programs. Policy scholars are still coming to terms with the nexus of formal and informal elements of the policy process, and the institutional and constructivist scholars are currently building on the insights of Barnard and others.

Stages Models of Public Policy

As the public policy subfield was developing, it relied heavily on case studies that permitted holistic examination of a single policy. These case studies suggested important generalizations about the policy process that extended the focus of policy scholars to include the examination of the political and administrative processes that preceded and followed formal adoption of policy. Very early in the subfield, a stages heuristic became the dominant model. The stages model typically identified agenda setting, policy formulation, policy adoption, policy implementation, and policy evaluation as the sequential processes of the model (Anderson, 1975; Jones, 1970). Theoretical case studies of the stages contributed to increased understanding of the policy process, especially in the areas of agenda setting, implementation, and evaluation, which had previously received less attention from political scientists than policy formulation and adoption. As useful as the stages heuristic was in organizing the policy subfield, it fell prey to intense criticism for a variety of reasons: no causal theory, insufficient research guidance, too little multistage research, insufficient hypothesis generation, imprecise prediction, and too much linearity (Sabatier, 1991). However, it should

be noted that even the scholars who were most closely associated with developing the model clearly indicated that the stages often were not distinct in actual practice and their order and characteristics could be quite varied (Anderson; Jones).

More than four decades later, the stages heuristic continues to anchor a substantial amount of policy research. The processes involved in getting the attention of the government, building coalitions of support, navigating the formal processes of policy adoption, crafting implementation, and modifying the policy over time continue to be essential elements of policy study even for those who are employing other, newer approaches.

Agenda Setting

Studies of agenda setting have tried to discern why some issues are given serious attention by government and others are not. Even among those that do receive serious attention, the question arises as to why some issues move quickly to reach agenda status and others take much longer. Thus, policy scholars sought to delineate the circumstances that make it more or less likely for a problem to be recognized and attended to by public officials.

Cobb and Elder (1975) observed the variation in the ability of groups to gain access to public officials and argued that this access influenced whether an issue was elevated to the formal agenda of a governmental entity. By the time public policy became a serious subfield, the importance of differential access to public officials had already been explored by the work of Schattschneider (1960). His work drew attention to the uneven participation in governmental decisions. His findings, that business interests and upper classes dominated public policy participation, flew in the face of the traditional pluralists' claims of policy openness. It also elevated examination of how governmental leaders moved some issues onto the agenda and blocked others. Two years later, another pivotal study added to this point by developing the concept of "non-decisions" and arguing that the blocking of certain issues from advancing onto the agenda is an important type of policy power that needs to be studied even if it is difficult to observe (Bachrach & Baratz, 1962). Bachrach and Baratz assert that disadvantaged groups are less likely to demand change when existing policies benefit widely accepted and powerful interests.

How a problem is defined affects whether and how public officials address it. Schattschneider (1960) was a very early voice in discussing this dimension of the policy process in his delineation of public versus private problems. To gain legitimacy and thereby earn a spot on the agenda, issues needed to be defined as public problems. When issues were judged to be inappropriate for governmental attention, there was little chance that the issue would move beyond a private issue (Eyestone, 1978).

Problem-definition research also placed perception and belief systems within the study of policy. Symbolic interactionists in sociology made the foundational contribution to this approach, arguing that human beings act on the basis of meanings they attach to things rather than on factual, objective definitions. This emphasis has found new energy in work by Stone (1988) and more recent scholars using the constructivist approach (Schneider & Sidney, 2009).

The multiple streams model articulated by Kingdon (1984) is sometimes given status as an independent approach to policy studies. It examines three separate streams of activity: problems, policies, and politics. These three dimensions of the policy process progress relatively independently of one another but occasionally couple, usually as a result of a policy entrepreneur who senses the opportunity to connect a problem with a policy proposal. When the political timing is right, the entrepreneur will push the coupled problem and solution through the policy window. Speed is important since the policy window of opportunity usually does not stay open long. Thus, agenda setting is intimately tied to an available solution and a political opportunity. This model is an excellent example of how public administration has contributed to the policy subfield since it is an adaptation of the so-called garbage can model advanced by three public administration scholars in their analysis of organizational processes and choice (Cohen, March, & Olsen, 1972).

Kingdon (1984) heightened the importance of politically capable individuals who he termed *policy entrepreneurs*. Policy entrepreneurs are opportunists who take advantage of crises or unanticipated events to push their policy proposals. Or they might take advantage of existing coalitions and work to soften up key policy communities. Another strategy is to modify the definition of the problem in order to take advantage of potential support to successfully push the proposal through the window. Ultimately, the entrepreneur is most successful when he or she is able to take advantage of political events or the national mood to promote a good idea that is technically feasible, consistent with core values, fiscally tolerable, and politically acceptable. However, there are many instances where issues fail to achieve or retain agenda status. Reasons for this include financial cost, lack of acceptance by the public or policy elites, opposition of powerful interests, and dominance of other issues (Kingdon).

Policy Formulation and Adoption

Policy formulation focuses on the drafting and consideration of proposals by an attentive policy community while policy adoption refers to the passage of the proposal through the formal institutions that have the authority to adopt the policy. This part of the larger policy cycle received the most scholarly attention prior to the development of

the policy subfield. It is closely associated with the institutions of government and is typically considered the key decision in the policy process.

In the multiple streams model, it is evident that the first three stages of the policy cycle—agenda setting, formulation, and adoption—take place more or less concurrently. Policy formulation itself focuses on the cognitive analysis and politics of designing a statute or other type of policy decision. The issue of rationality was a major focus of administration scholars in the middle of the 20th century. However, the debate between those arguing for rationality and against it continues. Today, rational choice theorists argue that a rational choice model is the best heuristic for studying, understanding, and predicting the outcomes of policy participants (Ostrom, 2007). The competing view is that scholars need to analyze language and political calculation to better understand the formulation and adoption processes (Allison, 1971; Halperin, 1974; Kingdon, 1984; Schneider & Sidney, 2009; Stone, 1988).

Many models incorporate both rational and political elements. Perhaps this is because the process itself incorporates both rational and political elements. Using a bounded rationality approach, policy analysts develop policy options, predict impacts, and oversee evaluation studies. But on the other side of the equation, political brokers approach the political landscape strategically, using data and analysis as well as other tactics to secure majorities at each stage in the adoption process.

One of the areas of greatest interest to policy scholars is the advocacy and policy design role of executive-legislative-interest group networks. Early scholars of political science advanced a number of models related to interest groups. One of the most noteworthy was David Truman (1951), who concluded that significant interests would organize to influence policies of concern to them. According to Truman and other pluralists, government's role was often one of facilitating and legitimating group compromise. In addition, pluralists suggested that executive agencies often participated in the bargaining, protecting their interests similar to any other group.

Additional models of executive-legislative-interest group relations preceded the development of the policy subfield and more have been created since. All of these models include groups inside and outside of government and portray coalition formation organizing around a specific policy interest. Numerous terms have been assigned to these models, including iron triangles, subgovernments, policy subsystems, issue networks, policy monopolies, and advocacy coalitions. Each of these versions of policy communities assumes the importance of relationships among actors interested in the policy issue. These models are not exclusively used in the American political context. They have been broadly applied to research in multiple political systems by scholars in many disciplines. In the context of the United States, however, these policy communities thrive in part as a result of the decentralized and fragmented

structure of its political institutions. Given the inability of the president and Congress as a whole to be informed and active participants on all issues, those issues that are not receiving widespread public attention and news coverage tend to be left to communities of people who have a deep interest in the specialized policy area.

In the 1960s, these communities were thought to be very stable relationships. In several case studies, the relationships were analyzed to be so stable that they were referred to as iron triangles (Cater, 1959). This conveys a pejorative view of the activity largely because the relationships among the governmental entities in the triangle and the interest groups were characterized as closed. New interests who wanted to communicate a competing set of demands and proposals were not acknowledged. Freeman (1955), in his studies of Indian affairs, showed how these policy subsystems brought stability to a policy area by bridging the executive and legislative branches with key interest groups. Each of these entities needed good relationships with the other two to enhance the likelihood of advancing its institutional and policy goals. This mutual dependence meant that once the subsystem emerged, its accommodations became the basis for determining which issues were placed on the agenda and which ones were not. These patterns of accommodation could be broken up by changes in committee leadership, widespread media attention to an issue in the policy area, or presidential interest. Once the visible interest in the policy topic waned, the subsystem political dynamic was likely to reemerge.

Hugh Heclo (1978) viewed the relationship between policy and administration as a vitally important one, especially as the role of government grew. He suggested that in the search for iron triangles, scholars sought to discover an exceptionally powerful and autonomous executive-legislative-interest group cluster of actors who dominated policy making by policy area. In the process, he believed that policy scholars ignored the more open and more commonly existing webs of people he referred to as issue networks. Issue networks are composed of those who are knowledgeable about the issue in terms of substantive knowledge as well as the history of its policy twists and turns. Policy knowledge is more heavily emphasized in this model, and it is the primary means through which additional participants can join this fluid web of relationships. In addition to being more fluid and episodic than subgovernments, issue networks, as shared-knowledge groups, have more points of view and conflicts than the iron triangle and subsystem frameworks. As aspects of the policy debate change, so do participants in the network. Thus, iron triangles and subsystems may still exist in some policy areas at some points in time, but the more typical pattern is one of a looser, increasingly complex kaleidoscope of policy (Heclo, 1978; Meier, 1985; Sabatier, 1991). For administrators and legislators, it provides a less stable and less predictable arrangement but one that permits greater maneuverability as well since these skilled

politicians have the potential to split, expand, and recombine the many segments of the issue network. The irony of the issue network is that, when compared with iron triangles, it expands the number of participants involved, thereby making policymakers contend with more conflict among multiple points of view. It also accepts that involvement in the policy process is contingent on a greater understanding of the increased complexities of the policy area. Heclo (1978) suggests that this pattern of knowledge-based participation may actually increase the cynicism of the general public as the gap between activists and the public expands.

Heclo's (1978) work was a pivotal change in the sub-field's orientation toward the nature and behavior of policy participants. Since Heclo, numerous scholars have articulated a variety of more diverse webs of policy participants, from Meier's (1985) regulatory subsystem to Sabatier's (1988) advocacy coalition framework. In the final analysis, though, whether one uses a looser, more episodic network or a more centralized and stable subsystem, the general public has virtually no role in any of the models. Given the level of specialized learning necessary to truly engage in the conversation among participants in the policy process, the general public's role is minimal unless the issue somehow ignites widespread interest.

Decision-Making Models and Policy Formulation and Adoption

In the early literature on decision making, rationality was elevated to a normative standard. Operations research during World War II contributed to an expectation for clear, measureable objectives, extensive research, and, ideally, evaluation of choices based on evidence. The rational-comprehensive model that emerged ideally required the clarification and prioritizing of objectives, followed by identifying a comprehensive range of options for achieving the ranked objectives, analyzing the capacity of each option to maximize ranked objectives, and choosing the alternative that best achieves the objective at the least cost. This model has continued to retain its normative appeal, but scholars have repeatedly shown that actual decisions are not made this way. Lindblom (1959) argued that the model breaks down in its first step since people find it exceedingly difficult to agree on the relative priorities of values and goals. In the public realm, policymakers must wrestle with conflicting values among the various participants involved. Even the thinking of a single participant often includes conflicting and unresolved priorities among values. Thus, these values are not usually clarified and rank ordered prior to designing a policy. Therefore, policies often embody conflicting values. Lindblom suggested that most policy is made following a process of successive limited comparisons. Analysis is truncated to a few feasible alternatives that are incrementally different from existing practice, and choice is made based on which option

receives consensual support. Pluralistic preferences and bargaining processes fit the incremental model well. Furthermore, this model can be expanded to a more intentional and strategic process for achieving substantial change since it is usually easier to achieve several successive incremental changes in policy over several years rather than attempting to secure support for major change in the first instance.

One of the most important contributors to the theories of decision making is Nobel prize winner Herbert Simon (1957). He argued that there are many reasons that rational-comprehensive models are not possible. First, information is lacking and people are not likely to be able to identify all possible alternatives. Second, the ability to accurately predict the outcomes of the many possible alternative choices is unlikely. And third, humans do not have the cognitive capacity to know and remember all that is required by the comprehensive model. Given these limitations, human beings accept what Simon terms *satisficing*. Under this approach to decision making, it is only necessary to find a solution that meets goals at an acceptable level rather than an optimal level.

Graham Allison's (1971) classic work analyzing the Cuban missile crisis shows how three different models of decision making each lead to very different explanations of the crisis. His rational actor model can be compared to Simon's bounded rationality model in that the goal is to choose the alternative that advances the national interest. The rational actor model assumes that nations function as centralized unitary actors where policy choices are made to maximize the national interest. Allison uses chess analogies to emphasize the strategic elements of the choice equation. Thus, the language of the model is of an optimizing effort, but the model recognizes the knowledge limitations and additional uncertainties that constrain choice. The advantages of this model are that it is simplified and stabilizes dramatically the information one needs to make choices. Morton Halperin (1974) emphasizes that given the number of bureaucratic departments and bureaucrats involved in decision making, it is difficult to confirm the unitary actor assumption of the model. The rational actor model assumes that all policymakers agree on the interpretation of the national interest. Halperin argued that this level of unity on what constitutes the national interest is rare in American history. One period followed World War II when the emergence of the cold war and the fear of communism solidified views of the national interest. This type of consensus also emerged for a short period after September 11, 2001.

Allison's (1971) second model examines governmental decision making as a function of the output of organizations. The organizational process model, which borrows heavily from public administration literature, argues that organizations serve as the primary actors in governmental decision making. Given their hierarchy and centralization, each organization functions in a unified manner alongside

other organizations. Under this model, analysts need to consider the outputs of multiple organizations. These outputs are the result of standard operating procedures, and organizational choices are in line with Lindblom's model of incrementalism. The best predictor for what organizations will produce in the future is to examine the status quo. Importantly, organizational decisions and behaviors are influenced by existing routines and values of the organization.

The third model Allison (1971) advanced was the bureaucratic politics or governmental politics model. This model is probably the one most closely attributed to Allison since it was the most original of the three. Under the governmental politics model, there are numerous individual participants who influence governmental choices and behaviors. Government decisions are more resultants rather than choices since decisions are a combination of actions by numerous participants in the process. Many of the participants become part of the process as a result of an organizational affiliation that they have, and they often take actions based on the values and objectives of their organization using the action channels their organizations provide them. As each of the participants takes actions advancing their personal and organizational interests, they may be involved in overt bargaining with other participants. Equally often, though, participants are inclined to take the actions permitted by their position in the system. The governmental decision is really the interaction and summation of all of these independent decisions rather than coordinated intention. Obviously, this model of decision making is not based on a single set of organizational or national values. Pluralistic values and actions dominate the process.

Just as public administration scholarship was important to the understanding of agenda setting, it also plays a central role in formulation and adoption processes. Both the subsystem and network approaches and the decision-making approaches rely heavily on public administration literature and concepts. The level of bureaucratic participation in policy making becomes even more the focus of policy making in the implementation of public policy.

Implementation

Implementation includes the administrative activities that convert a statute or other authoritative policy into a functioning program. Traditionally, implementation was characterized as a simple process of following the directives in the statute, administrative rule, executive order, or court ruling. Since early studies assumed that this took place without much delay or discrepancy from the policy's intent, relatively little attention was paid to this process prior to the emergence of public policy as a subfield.

One study that stands out for being well ahead of the development of the policy subfield was sociologist Phillip

Selznick's (1949) study of the Tennessee Valley Authority. This case study is one of the few studies that provided guidance to the complexity of implementation and the degree to which the contours of a policy could be significantly altered based on the decisions of implementers. Since his study focused on organizations as organic systems that adapt to their external environment, his findings created an awareness of how external forces cause policies adopted in Washington to be altered in the field. Selznick showed how organizations were dependent on local support and how local opposition from powerful groups generated adaptations in policy.

Selznick's (1949) study also challenged the traditional top-down view of implementation. Under the traditional top-down framework, bureaucrats function as instruments of the policymakers and respond to the command and control of those above them in a process similar to hierarchical lines of authority in a bureaucratic organization. However, scholars soon realized that implementation processes were much more complex and evolutionary than initially thought. Numerous studies showed that implementation was not faithful to the original plan for a variety of reasons. First and foremost, statutes and other formal policies often did not communicate with 100% clarity. There were gaps, overlaps, and contradictions. Thus, even if a bureaucrat was willing to function as a Weberian machine, it was typically the case that bureaucrats had to make significant decisions throughout the process of designing implementation. Even Woodrow Wilson (1887), whose work contributed to the policy-administration dichotomy framework, realized that all administrators operated with some discretion. Although the policy-administration dichotomy was successfully characterized by the public administration literature as unrealistic, policy scholars assumed the concept in their top-down implementation models until sufficient implementation case studies accumulated to convince most that the paradigm was inaccurate. Although there are still a variety of approaches to the study of implementation, most would agree that statutes are altered at least to some extent in the process of implementation. Policies are not self-implementing.

Another complexity in the implementation process that affects whether implemented programs align with statutes is that many national programs are carried out at the state and local levels, which dramatically increases the number of policy actors involved. This increase in the number of decision points creates huge difficulties for timely implementation and necessary communication and coordination. Pressman and Wildvasky (1973) masterfully show that as the number of participants and decision points increase, the likelihood of accurate implementation declines.

Two-way interaction models were a reaction to the inadequacies of the top-down models. Bardach (1977) and Lipsky (1979) posit that statutes and other authoritative decisions made by policy-makers interact with decisions of implementers to create the actual policy. As implementers work through the issues of how to carry out the

statute, they make choices that modify the policy. In addition, “street level” bureaucrats make use of discretionary authority as they engage in the day-to-day work of public policy. Accepting that implementation is an interactive process rather than a command and control hierarchy dictated by legislators focuses attention on the distance between a policy as originally designed or passed and the policy as implemented.

Majone and Wildavsky (1984) take the interactive model one step further in their work “Implementation as Evolution.” It posits that having the implementing agency modify the original policy design may actually produce beneficial results relative to the original goals of the statute. For example, if the original statute makes assumptions about cause and effect that are inaccurate, then having implementing bureaucrats modify the policy through implementation strategies would be a positive step as long as the changes they are making are in line with a commitment to the objectives and impacts the policy was supposed to produce. Both the implementation as evolution and the other two-way interaction approaches make it obvious how important it is to have agencies that are committed to the policy objectives.

Building on the work of all of these scholars, Mazmanian and Sabatier (1983) develop a synthetic conceptual framework to guide implementation studies as part of the larger policy process. They assert that it is necessary to examine the statute or other formal policy statement to assess its clarity of objectives, accuracy of causal assumptions, and effectiveness of decision rules provided to the implementing agencies. In addition, the researcher must appraise the sufficiency of financial resources, the specificity of the authority relationships, and formal access by outsiders. Variables beyond the statute and outside the implementing agency are also emphasized, including socioeconomic conditions, media attention, public and constituency group support, enthusiasm of higher level public authorities, and the commitment and skill of agency leaders. In addition to internal and external variables, there is also a recognition that implementation success is affected by the difficulty of the problem and the degree of change being sought.

Given the potential for significant adjustments to policy to occur as a result of implementation decisions, a diachronic approach of studying policies over a decade or more is recommended. Without a long-term view, the ability to understand the evolution of policy as it occurs in implementation will be incomplete.

Evaluation

Most stage models of public policy end with evaluation, the systematic assessment of the policy’s impact. Importantly, policies need to be examined for both intended and unintended consequences. Evaluations are

completed by the implementers themselves and by external policy analysts. Internal evaluations conducted by the implementing agency have the benefit of getting those who work with the policy on a day-to-day basis to recognize problems and propose solutions. Evaluations by outside analysts tend to bring external perspectives to the process and may possess increased legitimacy from the perspective of elected officials. One trend in the last several decades is for more ideologically affiliated think tanks to publish external policy analyses. These are usually provided with a particular point of view and thus do not necessarily offer the benefits of other external evaluations.

Distinctions should also be made between the evaluation of policy outputs and policy outcomes. Analysis of outputs draws attention to whether the administrative processes are in place through such measures as number and types of clients served. Outcomes refer to whether the policy is achieving desired results on policy goals as well as other unintended impacts.

For policy scholars, impact and outcome studies have received more interest than output studies. Scholars have also sought to understand how policy analysis leads to revisions, transformation, or termination of policies by elected officials. As could be seen in the previous discussion of implementation, revision of policy is often ongoing from the beginning of implementation decisions. However, most stage models have viewed evaluation as isolated from implementation and as more associated with the judgments of policy-impact recommendations for policy revisions. There is a tendency for this stage to look something like the feedback loop of the political systems model.

Systematic policy analysis grew as an important part of the policy process during the 1960s as government attempted to apply economic theory to policy making. Cost-benefit analysis, operations research, and various program performance measures incorporated in budgeting processes were primary tools of the effort. Thus, this part of the policy process is most closely associated with rational models of decision making. It was not unusual for program evaluation and periodic reporting to be statutorily required to assist in legislative oversight and budgeting processes. As the national government expanded the number of large domestic programs, along with detailed prescriptions and administration, collected evidence mounted that the programs were not achieving their lofty goals. Accountability continues as a common refrain today, but the capacity to systematically evaluate programs and to redesign them with effective performance measures has not been as easy or as successful as hoped. Once again, it is possible to see the tension between expectations for rational decision making as the vehicle for better policies and the reality that suggests there are significant human and organizational limitations to rationality. Even when systematic analysis is required, the uncertainty surrounding appropriate measures and the interpretation of results make evaluation as much politics as science.

Deborah Stone (1988), one of the leading critics of rational models of public policy, advanced a concept she termed *policy paradox*. She argued that politicians typically have dual goals: policy objectives and political objectives. Furthermore, she observed that analysis itself is political since it is rife with framing, definitions, and interpretations. Political participants frequently articulate an argument that on its face appears linear and rational but on closer examination appears constructed to achieve a political purpose. She offered the concept of political reasoning, rather than rational decision making, to understand the struggles of policy communities competing over which ideas, policy definitions, and corresponding solutions will prevail.

A closely related facet of the evaluation literature is the attention given to the use of knowledge generated through policy analysis. Generally, scholars have concluded that knowledge is not the most important dimension of policy decisions. Politics and the limitations of human and organizational capacity intervene (Simon, 1957). Even when decision makers seek information to help make policy, it is frequently questioned, interpreted in different ways, expensive, and incomplete. Although many write very pessimistically about the lack of use of policy knowledge, Carol Weiss (1977) argued that policy knowledge has been successfully used to identify problems, reconsider policy strategies, and provide an enlightenment function. Through the enlightenment process, information accumulates and causes policymakers to redefine the problem, retroactively make sense of why programs did not succeed, or readjust policy objectives to more realistic levels. Finally, policy knowledge is more likely to be used when it is timely and when participants see strategic benefit in doing so.

Alternative Models in the Study of Public Policy

In 1991, Paul Sabatier criticized the stages heuristic and challenged the field to develop better models. Sabatier and Jenkins-Smith (1993) put forward an alternative heuristic referred to as the advocacy coalition framework. They suggested that the stages model is too simplistic given the number of participants, institutions, and influences in the policy process. In their more complex model, they incorporate governmental and nongovernmental institutions, external conditions and events, and the multiples webs of these entities that align or compete with each other to influence policy. Furthermore, whereas the stages heuristic signals that policy follows a linear process, the advocacy coalition framework sees policies as nonlinear and rarely terminating. Multiple policies affecting the policy domain overlap and affect one another, as do policy implementations at multiple levels of government. Researchers who study a policy arena for several decades will observe policy adaptations and eventually significant change. This occurs because of learning on the part of the participant

coalitions that may compete over preferences based on different resources and belief systems. However, the advocacy coalition framework posits that the most substantial changes in policy are more likely to come from external events and conditions than from policy learning. To be able to witness this, scholars must be prepared to follow a policy area for several decades since advocacy coalitions tend to be relatively stable. Some scholars (Lester & Stewart, 2000) see the advocacy coalition framework as a development within evaluation research that could be incorporated within the stages model, but Sabatier and Jenkins-Smith (1993) claim that the model provides an important and enhanced alternative model.

Another model of note is the punctuated-equilibrium framework (Baumgartner & Jones, 1993). This model is similar to the advocacy coalition model in terms of the need to look at policy over long periods of time since most policy follows incremental change and relatively stable patterns over many years before incurring major change. Major punctuations occur in the context of changed beliefs among the policy community and often a new venue for implementation. Once implemented, equilibrium is reestablished, and the policy arena goes back to a lengthy period of stability. The punctuated-equilibrium model also draws attention to the importance of institutions since they tend to help ingrain the results of the dominant coalition and to resist pressures for change.

Elinor Ostrom (2007) and many others articulate a preference for a model of institutional rational choice. This model builds on the discipline's traditional emphasis on institutional structures and rules. Ostrom argued that institutions should be defined as rules, norms, and strategies that characterize entities with repeated processes. Rules dictate who has advantages in the pursuit of policy preferences and frequently determine who the major players are in the policy. This model also examines the hierarchical ordering of rules. Thus, constitutional rules influence options for players at other levels of policy choice all the way down to the operational, day-to-day decisions of policy actors. Rules can take the form of formal, written provisions or can be norms based on shared understandings by participants or guides that individuals develop to direct their own behaviors. These working rules help to provide stability in the midst of uncertainty. Once the study of levels and types of rules is achieved, the researcher can make probabilistic predictions using rational choice analysis. This model brings attention back to the formal versus informal distinction that was made by administration scholars many decades earlier.

Future Directions

One must ask whether any of the models will rise to a position of dominance in the near future. Given that scholars agree that the policy process is complex and

patterns of human behavior are varied, the search for a dominant, robust, parsimonious model is unlikely to be successful. This effort is also exacerbated by the number of contributing disciplines to the subfield. Frustration over this situation is also likely to continue. Perhaps Kenneth Meier (2009) captures it best when he complains that there are so many models, their ability to guide research is analogous to the interstate highway system's ability to guide a vacation. A more positive point of view may come from the realization that there is a lot of overlap among the various models. For example, the advocacy coalition framework, stages framework, punctuated equilibria model, and the institutional rational choice model all describe the importance of institutional structures, political brokers, external influences, and shared meanings.

Looking at policy processes more holistically and studying policies over the long term to incorporate evolution and change are recent trends in policy research that are likely to continue in the future. The role of beliefs and the processes leading to shared meanings also appear to be increasing in importance (Schneider & Ingram, 1997). It remains to be seen whether future studies integrate policy and public administration models. The literature of the two subfields has numerous parallels. Very few policies can be implemented without the significant involvement of an administrative system. The fact that the same policy can produce such different results across different implementing subunits (e.g., states, counties, or schools) suggests that agencies and individuals make a huge difference. Leadership styles and administrative cultures may offer vital insight to explaining the variation (Hicklin & Godwin, 2009; Robichau & Lynn, 2009).

Given the large number of models and participating disciplines, the methodologies employed in the subfield will continue to range from qualitative to quantitative. The quantitative methods employed in large-*N* comparative studies and some evaluation studies are likely to continue, as are the more qualitative methods of case studies and language analysis approaches. This is not necessarily a negative. As Easton (1969) argued in his call for a post-behavioral revolution, the important objective is to be relevant even if it means sacrificing the quantitative methods called for in the behavioral movement. Thus, a mixture of quantitative and qualitative measures is likely to continue to characterize the policy subfield.

References and Further Readings

- Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.
- Anderson, J. (1975). *Public policy making*. New York: Praeger.
- Bachrach, P. S., & Baratz, M. S. (1962). Two faces of power. *American Political Science Review*, 56, 947–952.
- Bardach, E. (1977). *The implementation game*. Cambridge: MIT Press.
- Barnard, C. I. (1938). *Functions of the executive*. Cambridge, MA: Harvard University Press.
- Baumgartner, B. D., & Jones, F. R. (1993). *Agendas and instability in American politics*. Chicago: University of Chicago Press.
- Cater, D. (1959). *The fourth branch of government*. Boston: Houghton Mifflin.
- Cobb, R. W., & Elder, C. D. (1975). Issue creation and agenda building. In J. Anderson (Ed.), *Cases in public policy making* (pp. 10–21). New York: Praeger.
- Cohen, M., March, J., & Olsen, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, 17, 1–25.
- Easton, D. (1953). *The political system*. New York: Knopf.
- Easton, D. (1969). The new revolution in political science. *American Political Science Review*, 63, 1051–1061.
- Eyestone, R. (1978). *From social issues to public policy*. New York: Wiley.
- Freeman, J. L. (1955). *The political process: Executive bureau legislative committee relations*. New York: Random House.
- Gulick, L., & Urwick, L. (Eds.). (1937). *Papers on the science of administration*. New York: Institute of Public Administration.
- Halperin, M. H. (1974). *Bureaucratic politics and foreign policy*. Washington, DC: Brookings Institution.
- Heclo, H. (1978). Issue networks and the executive establishment. In A. King (Ed.), *The new American political system* (pp. 87–124). Washington, DC: American Enterprise Institute.
- Hicklin, A., & Godwin, E. (2009). Agents of change: The role of public managers in public policy. *Policy Studies Journal*, 37(1), 13–20.
- Jones, C. O. (1970). *An introduction to the study of public policy*. Belmont, CA: Wadsworth.
- Kingdon, J. W. (1984). *Agendas, alternatives and public policies*. Boston: Little, Brown.
- Lester, J. P., & Stewart, J. (2000). *Public policy: An evolutionary approach*. Belmont, CA: Wadsworth.
- Lindblom, C. E. (1959). The science of muddling through. *Public Administration Review*, 19, 78–88.
- Lipsky, M. (1979). *Street level bureaucracy*. New York: Russell Sage Foundation.
- Majone, G., & Wildavsky, A. (1984). Implementation as evolution. In J. L. Pressman & A. Wildavsky (Eds.), *Implementation* (3rd ed., pp. 163–180). Berkeley: University of California Press.
- Mazmanian, D. H., & Sabatier, P. A. (1983). *Implementation and public policy*. Glenview, IL: Scott, Foresman.
- Meier, K. J. (1985). *Regulation: Politics, bureaucracy & economics*. New York: St. Martin's.
- Meier, K. J. (2009). Policy theory, policy theory everywhere: Ravings of a deranged policy scholar. *Policy Studies Journal*, 37(1), 5–12.
- Nelson, M. (1977, September). What's wrong with political science? *Washington Monthly*, 9(7), 12–20.
- Ostrom, E. (2007). Institutional rational choice. In P. A. Sabatier (Ed.), *Theories of the policy process* (pp. 21–64). Boulder, CO: Westview Press.

- Pressman, J. L., & Wildavsky, A. (1973). *Implementation*. Berkeley: University of California Press.
- Ranney, A. (1968). The study of policy content: A framework of choice. In A. Ranney (Ed.), *Political science and public policy* (pp. 3-21). Chicago: Markham.
- Rivlin, A. M. (1971). *Systematic thinking for social action*. Washington, DC: Brookings Institution.
- Robichau, R. W., & Lynn, L. E., Jr. (2009). The implementation of public policy: Still the missing link. *Policy Studies Journal*, 37(1), 21-36.
- Sabatier, P. A. (1988). An advocacy coalition framework of policy change and the role of policy oriented learning therein. *Policy Sciences*, 21(2), 129-168.
- Sabatier, P. A. (1991, June). Political science and public policy. *PS: Political Science & Politics*, 24(2), 143-147.
- Sabatier, P. A., & Jenkins Smith, H. C. (1993). Policy change and learning: An advocacy coalition approach. Boulder, CO: Westview Press.
- Schattschneider, E. E. (1960). *The semisovereign people*. New York: Holt, Rinehart & Winston.
- Schneider, A., & Ingram, H. (1997). Social constructions in the study of public policy. In J. A. Holstein & J. F. Gubrium (Eds.), *Handbook of constructionist research* (pp. 189-211). New York: Guilford Press.
- Schneider, A., & Sidney, M. (2009). What is next for policy design and social construction theory? *Policy Studies Journal*, 37, 103-120.
- Selznick, P. (1949). *TVA and the grass roots*. Berkeley: University of California Press.
- Simon, H. (1957). *Models of man: Social and rational*. New York: Wiley.
- Stone, D. A. (1988). *Policy paradox and political reason*. Glenview, IL: Scott, Foresman.
- Taylor, F. W. (1911). *Scientific management*. New York: Harper & Row.
- Truman, D. (1951). *The governmental process*. New York: Knopf.
- Weber, M. (1946). Bureaucracy. In H. H. Gerth & C. W. Mills (Eds. & Trans.), *From Max Weber: Essays in sociology*. New York: Oxford University Press.
- Weiss, C. H. (1977). Research for policy's sake: The enlightenment function of social research. *Policy Analysis*, 3, 531-546.
- White, L. D. (1926). *Introduction to the study of public administration*. New York: Macmillan.
- Wilson, W. (1887). The study of administration. *Political Science Quarterly*, 2(2), 197-222.

92

CAMPAIGNS

CHAPMAN RACKAWAY

Fort Hays State University

Political campaigns represent the core of representative democracy. To win an election, a candidate must earn the support of the general public. The quality of a democratic society can be easily linked to the quality of its election campaigns. Vigorously contested elections and widespread voter participation are two of the hallmarks of a strong democratic nation. Scholarship into campaigns therefore has significant implications for the quality of democracy.

As parties weakened from the beginning of the 20th century, candidates (and therefore campaigns) became much more important. Candidates have taken on the responsibility of organizing and funding their contests after more than a century of partisan control over campaign administration. In the intervening century, campaigns have become more varied in their approaches, structures, and strategies. The continuing evolution of campaigns and campaigning has created a rich area for academics to study.

However, campaigns have been the subject of serious scholarly attention for only the last 20 years. The bulk of academic work in election campaigns has focused on the resulting voting and citizen participation in those campaigns. The campaigns themselves—the organized efforts to motivate voters to support a particular candidate—had been largely ignored until the 1980s. Over the last 20-plus years, though, campaign scholarship has developed significantly.

One of the great difficulties in studying campaigns is the diversity of styles that exist at different levels of election. For instance, a contest for the U.S. Senate will be almost unrecognizable from the nationwide spread and focus of a presidential contest. A city commission election, which is tiny in comparison to either the senatorial or presidential level, will barely be noticeable in campaign activity compared with either of the other contests. Just because there is a contest for political office does not mean that all campaigns are the same. Some campaign elements, such as strategy, will be universal. But in almost every way, campaigns are beholden to the level of office sought. The diversity of levels and styles means that general theory building is much more difficult in the area of campaigns.

Theory

Scholarship of campaigns can be organized according to six distinct areas of the campaign: (1) level of the campaign, (2) organization and professionalism, (3) differences between genders in campaign style, (4) campaign advertising, (5) fund-raising and fund-raising-law compliance, and (6) campaign effects. Although some areas overlap, the field of campaign scholarship has focused on those six areas primarily.

The level at which the campaign is conducted makes a significant difference in how a campaign is administered.

A presidential campaign, with hundreds of staff, hundreds of millions of dollars to spend, and a 50-state approach, will be naturally rich in resources and highly competitive. As the size of the constituency decreases, so will most or all of the campaign elements such as money and professionalism. Campaign scholarship is therefore first and foremost subdivided by office sought.

The area that will likely show the greatest diversity of approach is in the organization and professionalism of a campaign. Some campaigns will be run by the candidate with the help of a select group of volunteers. As the campaign level gets higher, the professionalism level will as well. Campaigns will increase in both number of staff and the likelihood that the staff will be paid.

Questions persist as to the representativeness of candidates for office. Campaigns for office have been largely the province of men, but women have a different style of campaigning, and the presence of a female candidate significantly changes the dynamics of a campaign. As a result, there is a growing scholarship on the role of women in campaigns.

Campaign advertising has been studied extensively, from the effectiveness of televised ads to the reach of direct-mail advertising. As perhaps the most visible part of high-level campaigns, campaign advertising is perhaps the most studied element of a campaign.

Fund-raising is the most easily quantifiable element of a campaign and thus another area of intense scholarship. Campaign money has to be separated into two areas: the money taken in and the reporting to relevant compliance authorities. Campaigners have to raise money, and how they do their fund-raising has become an area of significant academic attention. The limits and disclosure requirements attached to campaign finance by both the states and the federal government are also areas of great interest to scholars in the field.

Finally, there is the question of whether campaigns actually affect the voting population at which they are targeted. A campaign is assumed to be the mechanism through which voters make their decisions on whom to vote for. However, if campaigns leave no measurable effect on the voters, then their value is in question.

Applications and Empirical Evidence

Level of Campaign

The presidential election is unique in American politics, the producer of constant campaign material for 18 months and longer. Sidney Blumenthal first termed the presidential contest the *permanent campaign* in 1982. Ornstein and Mann (2000) critique the constant campaigning, pointing out that once elected, political figures cannot disconnect from campaign mode and that policy is made more for the purposes of re-election than solving collective problems.

Presidential primaries are a unique entity in and of themselves and subject to their own scholarship. The primaries are held early in an election year, but in fact the process of campaigning begins more than 20 years earlier. In presidential primaries, one significant question relates to representativeness of voters. If a presidential primary features low and unrepresentative turnout, then the implications for a democratic society are dire. Unrepresentative party nominees who are less appealing to the general election electorate can mean lower voter turnout and dissatisfaction with government. Barbara Norrander (2000) finds that presidential primary voters are representative of the larger electorate; however, Norrander finds that presidential primary election campaigns are essentially strategic contests and that the number of candidates in the primary is a result of strategic analysis. Candidates decide to abandon their presidential nomination campaigns based on the gap in poll numbers between them and the front-runner. Momentum is vital in a presidential nomination campaign, and if a challenger for the nomination sees that another candidate is building momentum toward nomination, their candidacy is under serious threat. Norrander's findings are consistent with the idea that presidential candidates are strategic actors who pay close attention to their standings in the polls.

The timing of the 50 primary contests is also important to nomination hopefuls. Mayer and Busch (2004) point to the increasing compaction of the primary calendar as driving up costs, limiting information available to voters, and making it more difficult to campaign for candidates. The main point is a problem of federalism: The national parties have very few rules that prevent front-loading, and states seek to advantage themselves by moving earlier in the calendar to be more important to the nomination process and therefore get more resources and attention from the campaigns. The parties are the solution, according to Mayer and Busch, and they recommend a significant overhaul of the primary process that is led by the national parties themselves.

The strategy of a presidential campaign is largely dictated by polls, as Sigelman and Buell (2003) found. Using the presidential campaigns from 1960 to 2000, the authors found that campaigns chose to use an attack strategy when they were behind in the polls. When one ticket had a significant lead over the other, the trailing campaign was always likely to attack. When the campaign in question held the lead in polls, any attacking was led by the vice-presidential nominee. Closely contested races, however, left strategic uncertainty for the campaigns with a resulting absence of any pattern. Sigelman and Buell take the fact that both parties' nominees focus on the same issues to develop an alternative theory of presidential issue campaigning, one of issue convergence. Between 1960 and 2000, the authors show a striking unanimity in the issue appeals by candidates of both parties. Sigelman and Buell point out that democratic dialog is lost when campaigns take such a homogenous view of issue positions.

Presidential campaigns must choose their issues and positions strategically as well. John Petrocik, William Benoit, and Glenn Hansen (2003) have developed the idea of issue ownership in campaigns: that one party has a reputation as being more trustworthy on an issue than the other, giving that party's nominee an advantage. Generally, Petrocik and his colleagues find that Republican-owned issues such as defense and taxes get significant attention from both Democrats and Republicans, which in turn provided Republican candidates for office an advantage in campaigning between 1952 and 2000.

Presidential campaigns constantly provide new content for the media and voters to consume. A presidential campaign is omnipresent during a campaign season, but lower level races display different characteristics. For instance, elections for the U.S. Senate vary greatly between state and cycle. Some campaigns may be as intense as a presidential campaign, with the candidates close in polls and daily messages coming from each candidate's camp. Others may have much less content communicated to voters and less apparent competitiveness. A 1997 study showed that intense campaigns—close contests with regular activity—encouraged voters to think more of their own ideology and the policy implications of electing either candidate. Voters still used traditional cues like political party affiliation and perceived presidential performance on the economy (Kahn & Kenney, 1997). In other words, close contests with lots of activity leads to voters deciding on their votes with more sophistication.

In campaigns for the U.S. House of Representatives, the barriers to entry are often more important than the campaign itself. District partisanship and incumbency can have a suppressive effect on potential candidates. Strategically minded candidates survey the political landscape before entering a contest, and a strong incumbent or a district where the party identification of voters may make their campaign tougher may become the reason a strong contender does not enter the race (Stone, Maisel, & Maestas, 2004).

Furthermore, congressional races are beholden to the top-of-the-ticket contests. Scholars have long established that public approval of the president matters to the electoral fortunes of congressional candidates, but the real determinant of congressional success in light of their presidential support is their own voting record. Even more than partisanship, a candidate's public support for the president is the main determinant of public attitudes toward the congressional candidate's presidential support. In effect, a member of Congress's vote relative to the president's position is key in garnering support of the constituents in their district (Gronke, Koch, & Wilson, 2003).

The literature on state-level campaigns is sparser than research on federal-level campaigns. State legislative campaigns are less competitive than higher level races, with more uncontested races and larger margins of victory than federal contests. The length of a state legislator's

term—full-time or professional legislatures—and having a traditional single-member district substantially increases advantages for incumbents (Carey, Niemi, & Powell, 2000).

Findings on state legislative races indicate district characteristics are strong predictors of challenger emergence against an incumbent lawmaker. Incumbents can prevent strong challengers through constituent responsiveness on the job, but responsiveness has only a minor impact on the vote (Hogan, 2004).

One significant area of note regarding state legislative races is the idea of those contests as a stepping-stone to higher office. Strategic ambition by state legislators drives them to pay more attention and be more responsive to constituent concerns (Maestas, 2003).

The strategic context of any campaign is driven primarily by the level at which the candidate is competing. So is the level of professionalism and use of advanced techniques. Early political campaigns at all levels were amateur-driven affairs. Even presidential campaigns featured a largely amateur staff. However, as candidates emerged as the primary organizing force in campaigns, professionalism followed.

Image consultants and media designers gradually became part of the campaign process. The presidential-level campaign saw the greatest reliance on professionals, but by the 1970s, senatorial races were commonly run by professionals. Paul Herrnson (1992) began chronicling the level of professionalism in congressional campaigns through a regular survey of candidates in the 1990s. In the decade-plus since, Herrnson's research has shown that congressional races are highly professional. Campaign managers are more experienced, more likely to be hired consultants or professional staff, and a greater number of campaign roles are filled by paid professionals instead of volunteers.

Organizing a professional campaign for Congress paid off in a variety of ways, most especially in the area of fund-raising. Herrnson (1992) showed in other work that a professionally managed campaign led to more fund-raising success for the candidate. Indeed, professionalism is one of the aspects of campaigning that has led to the increased spending associated with election contests. Professionals by their nature expect pay for their services, and staff salaries increase the need to fund-raise for campaigns. Professional staff in a campaign also know how to conduct fund-raising programs, so the more professional a campaign, the more likely that organization will bring more money in. Professionalism and fund-raising can be seen as a set piece, two campaign elements that go hand in hand.

The expansion of political consultants into the variety of campaign roles was chronicled by Herrnson's (1986, 1988, 1992) research. Herrnson points to an expanding slate of campaign roles, including professional fund-raisers, media consultants, and strategists. Campaigns, especially

at the congressional level and above, are now highly professional exercises. Burton and Shea (2003) show that not only have roles become professionalized, but also campaigns themselves follow a more established pattern as experienced campaign hands develop playbooks of methods and strategies that campaigns follow.

Professionalism in campaigns has its consequences, and one of those is a greater tendency toward attacks on the opposing candidate. Professional campaign consultants are commonly thought to move campaigns toward a more negative orientation. Other literature exists showing the incentive for a campaign to attack, especially being behind in the polls. An environment with polls and ample fund-raising is necessary to strategically determine the need to attack, and that means having political consultants. In other words, more professionalism means more strategy, and the strategic imperative is much more likely to lead to attacking campaign tactics than in nonprofessional campaigns.

The strategic and professional development of congressional campaigns has led to a new theoretical approach to campaigns. Burton and Shea's (2003) conception couples a desire to win with a commitment to understand everything possible about the voters of a given district, campaign laws, folkways, and issues. A successful and strategic campaigner, in Burton and Shea's view, works backward through a plan. Campaign planning is one of the vital parts of any electoral contest, including understanding past voting patterns, budgeting and media planning, potential opposition strategies, and issues to address. Professional campaigns are often characterized by the presence of a well-written and carefully constructed campaign plan.

One important driver in the increasing professionalization of congressional campaigns has been the resurgence of political parties. Herrnson (1986) countered theories of partisan decline by pointing out that national political party organizations were heavily involved in the administration of congressional races. In areas that require technical expertise and research, such as voter targeting, fund-raising, polling, and campaign finance law compliance, Herrnson found that political parties were highly involved in competitive campaigns. Most important, Herrnson points out that the parties are highly selective in which races they choose to devote resources to. The process of targeting helps political parties maximize the number of party members they have in Congress. By choosing races that are competitive, with the best chances of their party to win, the party committees like the National Republican Congressional Committee and the Democratic Congressional Campaign Committee ensure they have the best chance of earning a majority of seats in the legislature. The parties look for a number of important factors, such as the candidate's early fund-raising success, their name recognition among people in the district, and their experience campaigning. The parties then devote

significant resources including independent ads, campaign schools, advice on consultants to hire, and fund-raising assistance (Herrnson, 1988).

Herrnson's work is core in the area of campaign scholarship because of his understanding of the role of political parties. Throughout the 1970s and 1980s, party scholarship focused on the shrinking identification with parties in the electorate. However, through the 1980s and beyond, a cadre of scholars, including Herrnson, pointed to efforts by the party organizations that transformed how political campaigns were conducted. Campaigns, under the aegis of ambitious and organized political party campaign committees, became more structured and professional.

Parties helped make races more competitive by recruiting the best possible candidates they could. The political party organizations developed recruitment coordinators, usually elected members of the chamber, who networked with party leaders throughout the country to identify potential candidates and then begin grooming them for later runs for office by encouraging them to run for lower level offices. Kolodny and Dwyre (1998) described an acceleration of party efforts to orchestrate individual congressional campaigns in the 1990s. Beginning as an effort to rebuild political party organizations, the professionalization effort quickly evolved into a larger-scale effort that completely changed the culture of campaigning for office.

Congressional campaigns transformed, and other levels of campaigning followed suit. Abbe and Herrnson (2003) found a growing professionalization of state races, but only in selected states. A state campaign is much more likely to raise more money, hire consultants, and be more strategic when the legislature is professional and the state is close to Washington, D.C. Daniel Shea (1995) also showed that state-level politics was following in the same professional model that national politics had already undergone. Political party organizations in the states began to focus on recruiting, targeting resources, and professionalizing state legislative campaigns.

The effectiveness of partisan campaign involvement has been questioned in recent work. Regarding the efficiency of party contributions to congressional races, a significant disparity exists in the races money is distributed to. The parties simply try to find the best possible challenger and open seat candidates and then distribute money and other resources to them. The parties also act as liaisons between the campaigns and Political Action Committees (PACs) to further extend their fund-raising reach. Incumbents, particularly those in very safe districts, are expected to contribute back to the party committees for money to redistribute to seats that are either endangered or potentially picked up. The result is a rethinking of the political party as a coordinating entity. Schlesinger (1996) asserts that the new party reality is one of a primarily electioneering organization that acts as the

nerve center of a campaigning network that includes the candidates, consultants, PACs, and media.

Differences in Gender

All things being equal, female candidates should be just as successful in running for office as their male counterparts. From the advent of women's suffrage into the 1980s, though, female candidates were few, and elected female officials were even fewer. Female candidates are therefore an important area of scholarship as political scientists today attempt to discover why female candidates and representatives lag behind.

In the 1970s, there were only 13 women in the 435-member U.S. House, a single woman in the Senate, and no female governors. By 2008, 16 women served in the U.S. Senate, and 16 women were governors. More than 50 women were U.S. representatives, and almost a quarter of state legislators throughout the country were women. The disadvantage that female candidates were assumed to have had apparently disappeared.

A number of scholars find that when female candidates run, they win (Seltzer, Newman, & Leighton, 1997). Male candidates significantly outnumber female candidates, though, so the percentages of successes are not as obvious as the raw numbers of victories for men compared with women when running for office. At all levels of campaigns, by the 1990s the assumed disadvantage that female candidates had was not evident in office-seeking success.

For female candidates, 1992 was a breakthrough year. Commonly referred to as the Year of the Woman because of the significant number of women who ran for office and won, 1992 has received significant attention because of the potential insight for prospective female office seekers. A shift toward domestic issues after the cold war, recruitment efforts by the national party committees, and the rise of female-focused political groups such as EMILY's List contributed to the female-candidate boom in 1992.

Taking 1992 as a transformative year, then, lets us assume that being a female candidate is no longer (if ever) a detriment to successful campaigning. Accepting that assumption, a female candidate's gender can either be a nonfactor in a campaign or an asset. For the most part, being a female candidate today is an asset. One study shows that female candidates who focus on issues closely aligned with the wishes of their female constituents see their gender become an asset. The authors dispute the idea that the common stereotype of women as weaker candidates is outdated (Herrnson, Lay, & Stokes, 2003).

For some voters, gender is a cue in their information search. In a low-information environment, the simpler the cue, the better. The gender of a candidate is a very relevant and obvious cue. As a result, women tend to support female candidates, and gender becomes a highly significant informational cue for women voters. However, sometimes informational cues conflict. The advantage for female

candidates appears to only work for Democratic candidates. Since women tend to vote Democratic, the cues of party and gender reinforce each other. Since the cues of female and Republican conflict, the advantage that female Republican candidates have is much less than their Democratic counterparts.

Being a woman is not enough, though, to ensure a successful campaign (Atkeson, 2003). Many female candidates prior to 1992 were not part of the movement to professionalize campaigns, so their efforts came across to voters as amateurish and not serious enough. Atkeson shows that a candidate must be taken seriously enough by the party, the press, and outside interests to be considered as having a real chance to win. Without the perception that a female candidate can be competitive, just being a woman is not enough to ensure success or support by women.

Atkeson (2003) shows that stereotypes and assumptions about women persist in election campaigns. Democratic and Republican female candidates are not created equal. For Democratic female candidates, voters perceive them as being extreme in their ideologies. As a result, their Republican opponents are generally more successful in general election campaigns. For Republican female candidates, voters see them as closer to the middle of the ideological spectrum, improving their chances at winning. The most in-depth work on the ascension of female candidates is by Sanbonmatsu (2002), who presents two important findings. First, Sanbonmatsu points out that the social class differences between the parties means that female candidates have an easier time emerging in the Democratic Party. Second, her study shows that women are less likely to self-select in their candidacies, and so they become more reliant on party recruitment. Therefore, states that have more developed party organizations and recruitment mechanisms are more likely to recruit female candidates.

Sanbonmatsu's (2002) work is reinforced by Fox and Lawless's (2004) research. In the Fox and Lawless piece, women show generally lower ambition to run for political office than men. Again, women tend to not self-start in campaigns compared to male potential candidates. Two factors depress female candidacies: a perception of self by the potential candidate that she is not as qualified to run for office as others and a lack of recruitment efforts by others. Although female candidates might have replaced their disadvantages as candidates with advantages, the greatest deterrence to female success in campaigns is the initial decision to enter the race.

Campaign Advertising

The field of campaign advertising is an area rich in research. Televised campaign ads have become ubiquitous since their introduction with Dwight Eisenhower's "Ike for President" animated advertisement in 1952. Over time, campaign advertising has advanced well beyond the jingles

and slogans that characterized early ads. Consistent with other technologies, strategy, and professionalism, lower-level races have also embraced televised advertising as a method of reaching voters during a campaign.

Trent and Friedenbergs (2008) vital work on political communication relates the entirety of political communication to the need to campaign. For voters, campaign-related communication gives them the most concise and concentrated amount of political information possible. Trent and Friedenbergs then see all forms of campaign communication as beneficial and essential to a democratic society. More a history of campaign outreach, including debates as well as advertising, Trent and Friedenbergs do not delve into how effective ads are generally.

Early work on negativity suggested that the effect of negative campaign ads was severely limited. Negative political advertising evokes negative feelings toward the targeted opponent and also the sponsoring candidate. Voters seem to look at a negative strategy in blanket terms. Rather than simply accepting the attacks levied in a negative ad against a candidate, the voter who consumes an ad thinks negatively about both candidates, and therefore the likelihood is reduced that he or she will strongly support either candidate. Implied in the research is that voters exposed to a significant amount of negative ads may be less likely to vote.

A major breakthrough in the study of campaign advertising came in 1991, with the first book dedicated to negative campaign advertising. Johnson-Cartee and Copeland developed a dichotomy of negative advertising, claiming that attack ads find fault with the political positions or record of the opponent and also his or her individual character or personality. Much of the negativity assumed to be held by the public, the authors show, relates to the personal ads and not the policy ads. Johnson-Cartee and Copeland showed that at least some subtle critical thinking was being done by voters as they viewed the ads in question.

Indeed, many point to campaign advertising as a reason for the decline in elements of civic engagedness like beliefs in efficacy of government and voter turnout. Ansolabehere and Iyengar (1997) broke ground in the effects of negative campaign ads and, using experimental methods, exposed voters to carefully designed ads to measure the differences between policy attacks, personal attacks, and nonattack advertising. The study showed that voters who were exposed to negative political campaign ads were 5% less likely to turn out and vote. The primary reason that voters report less likelihood to vote is that they feel disengaged from and uncomfortable with the political process. The implication and assumption in the work of Ansolabehere and Iyengar is that voters are turned off by negativity in general. In a later book, the authors replicate and repeat the findings from their original work. Ansolabehere and Iyengar firmly established one side of the debate on campaign advertising. Negativity as a

depresser of civic engagement and voter turnout persists today as one theoretical orientation toward campaign advertising.

However, other research shows that campaigns should use negative campaign advertising because of its effectiveness with voters. Recent scholarship shows that political advertising contains significant informational and emotional content, contributing information, engagement, and participation. Using the 2000 election as a case study, one article clearly shows that citizens who are exposed to campaign advertising develop higher levels of interest in the election, think more about the candidates running, are more familiar with the candidates, and are more likely to vote. The most important finding of their research is that the beneficial effects of campaign advertising emerge mostly among those with the low levels of political interest and information prior to the beginning of the campaign (Freedman, Franz, & Goldstein, 2004).

More recent research has continued to dispute the idea that negative advertising suppresses engagement in campaigns and suggest they are in fact a vital part of the information search that voters involve themselves in. Using multiple statistical methods on data from American National Elections Studies over 42 years, Finkel and Geer (1998) showed that the reason for lower voter turnout and engagedness could be explained better by other factors than exposure to negative advertising.

The most important rejoinder to Ansolabehere and Iyengar's work comes from Wattenberg and Brians (1999), who also used American National Elections Studies data and other sources, showing that candidates who sponsor negative ads do not suffer a loss of support and votes from respondents who recalled seeing that campaign's negative ads.

For two decades, political scientists have debated the role of campaign advertising and its importance to the candidates who sponsor them. Two schools have emerged: one arguing that negative ads suppress turnout and lead to backlash against the sponsor and another that says the effect of exposure to negative ads is minimal and overpowered by other elements of the campaign.

Fund-Raising and Compliance

Another aggressively studied element of the campaign is fund-raising. Jesse Unruh, former speaker of the California State Assembly, coined the phrase "Money is the mother's milk of politics" (Cannon, 1969, p. 99). The phrase, first uttered in the 1960s, is still true today and is ever truer. As campaigns have become more professionalized, they have required even more money for consultant salaries, polling voters, developing and placing advertising, direct mail, and other strategic elements of campaigning.

Frank Sorauf (1994) published some of the early works on campaign finance. After the Federal Election

Campaign Act limited campaign money and mandated disclosure of donations and spending, data became available to analyze. Sorauf tracked the amounts of money donated to campaigns in aggregate as well as the spending by those campaigns. Sorauf in particular shed light on the common assumption that elected officials who accept donations from interested parties are bought and sold by those interests. Sorauf points out, in a theme that continues throughout most of the subsequent literature on the topic, that the assumption is based on correlation and not cause. In other words, candidates accept money from interests who are aligned with their preferences, but it does not mean that those groups are imposing their preferences on the elected officials to whom they donated.

The demand for money has led to a development of fund-raising styles, especially the personal touch as described by Francia (2003). Fund-raising committees and personal appeals by the candidate are the most successful elements of fund-raising by campaigns, using emotional language of threats by opponents and the need to fight against them.

At the presidential level, there are two primary methods of soliciting funds: direct mail appeals to voters and the use of interpersonal networks to bundle contributions. Since that work, though, the Internet has allowed for direct fund-raising online. Indeed, the Obama presidential campaign of 2008 used online fund-raising to dramatic effect, raising very small donations from a wide variety of voters. The small-donation model, dependent on social networks, has been established as a key method of successful fund-raising (Cho & Gimpel, 2007).

Greater attention has been paid to preprimary-phase fund-raising in recent years. One study focused on preprimary-fund-raising success using professionalism of the campaign, length of candidacy, and the candidate's poll performance. The author found that more money spent on fund-raising early in the preprimary phase, perceptions of candidate competitiveness, and poll results all increase a campaign's ability to successfully fund-raise (Goff, 2007).

Some scholars turn conventional wisdom on its head, actually positing the question of why so little money flows into campaign coffers. Reinforcing the initial findings of Sorauf, the work finds that small sums dominate campaign fund-raising and that the money does not have nearly the negative affect commonly assumed. Ansolabehere, de Figueiredo, and Snyder (2003) find that interested groups give little and get little but do so because of the importance of the marginal success they may have. The access to elected officials they get and the subsequent opportunity to influence legislators is enough to provide the groups an incentive to give just a bit. Additionally, the goals of the groups are more to affect elections and ensure friendly legislators are in office, knowing their policy success is not guaranteed.

Campaign Effects

The final question to ask is whether all of these campaign elements have any effect on voters. Early work on the effects of campaigns said that candidate outreach did little to nothing to add supporters. Most voters, the research showed, had made up their minds based on partisanship and approval of the incumbent. Finkel (1993) found that campaigns serve primarily to mobilize existing supporters, activating that support rather than converting opponents or swaying the undecided.

If there is a measurable and notable effect from campaigns, it comes in the form of voter mobilization. Recent research has focused on the efforts of campaigns and affiliated groups to get voters out to the polls. Local and county party organizations that are well organized and well led are more aggressive about mobilizing voters and therefore more successful (Beck, Dalton, Haynes, & Huckfeldt, 1997). Outside groups such as interests and PACs have begun interacting with parties more since the Bipartisan Campaign Reform Act of 2002 banned the practice of soft-money donations to parties from groups (Magleby, Monson, & Patterson, 2007).

Gerber (2005) and Green (2003) have led the study of effectiveness of campaign activities throughout the last decade. Their work has shown consistently that the personal touch, so vital in fund-raising, is also necessary to effectively mobilize supporters. Direct mail has a slight positive effect on mobilization, as do phone calls. Face-to-face campaigning, though, has the greatest effect on mobilization, and that effect is most pronounced among younger voters.

Policy Implications

The First Amendment to the Constitution provides a significant degree of protection to political speech, so regulations impact the campaign environment. Campaign finance is the most notable area of legal limits on campaign activity, and those regulations are common. The Bipartisan Campaign Reform Act (BCRA) of 2002 provides an example of the interaction of policy and campaigning. The BCRA banned unlimited soft money contributions from interested companies and groups to political parties for campaigning and had two significant effects on campaigns: First, they midwived the birth of 527 organizations, and they also led those groups to change from donors to mobilizers.

Reversing the causal chain, campaigns can have an impact on public policy. However, as the research on campaign finance shows, the money that flows into campaigns does not buy the votes or support of elected officials. The money buys only access and the opportunity to occasionally influence policy. One can say affirmatively that the common assumption about interested money buying and selling legislators is a myth easily dispelled by volumes of research on the effect of campaign money.

Future Directions

As the campaign scholarship field expands, a number of areas will emerge in the research. State-level campaigns are the most obvious area of greater research. Since federal campaigns have seen a greater unanimity of style and strategy, the diversity of political cultures among the states provides a rich field of opportunity to study the professionalism, fund-raising, and strategic elements of those campaigns with the same analytical attention paid to the national campaigns.

As the Internet supplants television as the dominant media that voters consume, the effects of web video as a campaign element should emerge as an important field of study as well. Indeed, a campaign can conduct many of the traditional elements of its outreach over the Internet, so fund-raising, volunteer mobilization through social media, and advertising content can all be studied from the new frame of Internet campaigning.

Finally, the effects of campaign activities on voters will continue to develop. Experimental models and survey research will likely continue to be the method through which scholars examine the effectiveness of campaigns, but more models will emerge and more data will show if the initial findings on mobilization are robust.

Conclusion

For a long time, campaigns were treated as a given, and scholars paid more attention to voting behavior as the only important element of a campaign. However, the last 20 years have seen an emerging scholarship on campaign elements. Professionalism has dominated federal-level campaigns, whether for the presidency or Congress. The professionalism has led to more strategic thinking on the parts of all campaigners and a more consistent style of campaigning.

The use of advertising has also increased and become nearly ubiquitous. The effects of campaign advertising, particularly attack ads, is an area of great dispute. Two schools of thought have emerged about the effectiveness of campaign advertising, one arguing it depresses voter turnout. The competing school claims negative ads provide useful information to voters and are a strong contribution to democratic debate.

Women as candidates, long assumed to have significant disadvantages, have emerged as overcoming the biases and perceived weaknesses that hurt their electoral chances in years past. Campaign fund-raising is closely linked with professionalism and running ads and has increased over the last few decades, but research points out that candidates are not beholden to the interests from which they fund-raise. Finally, campaigns are not the sole force that drives votes. Many voters have their minds made up prior to the campaign's beginning. However, campaigns can activate soft

supporters and do convert the undecided. In close races, therefore, a campaign's activities are a vital contribution to the democratic process.

References and Further Readings

- Abbe, O. G., & Herrmson, P. S. (2003). Campaign professionalism in state legislative elections. *State Politics & Policy Quarterly*, 3(3), 233-256.
- Adkins, R. E., & Dowdle, A. J. (2002). The money primary: What influences the outcome of pre primary presidential nomination fund raising? *Presidential Studies Quarterly*, 32(2), 256-275.
- Ansolabehere, S., de Figueiredo, J. M., & Snyder, J. M., Jr. (2003). Why is there so little money in U.S. politics? *Journal of Economic Perspectives*, 17(1):105-130.
- Ansolabehere, S., & Iyengar, S. (1997). *Going negative: How attack ads shrink and polarize the electorate*. New York: Free Press.
- Atkeson, L. R. (2003). Not all cues are created equal: The conditional impact of female candidates on political engagement. *Journal of Politics*, 65:1040-1061.
- Beck, P. A., Dalton, R. J., Haynes, A. A., & Huckfeldt, R. (1997). Presidential campaigning at the grass roots. *Journal of Politics*, 59(4), 1264-1275.
- Burton, M. J., & Shea, D. M. (2003). *Campaign mode: Strategic vision in Congressional elections*. Lanham, MD: Rowman & Littlefield.
- Cannon, L. (1969). *Ronnie and Jesse*. New York: Doubleday.
- Carey, J. M., Niemi, R. G., & Powell, L. W. (2000). Incumbency and the probability of reelection in state legislative elections. *Journal of Politics*, 62(3), 671-700.
- Cho, W. K. T., & Gimpel, J. G. (2007). Prospecting for (campaign) gold. *American Journal of Political Science*, 51(2), 255-268.
- Finkel, S. E. (1993). Reexamining the "minimal effects" model in recent presidential campaigns. *Journal of Politics*, 55(1), 1-21.
- Finkel, S. E., & Geer, J. G. (1998). A spot check: Casting doubt on the demobilizing effect of attack advertising. *American Journal of Political Science*, 42(2), 573-595.
- Fox, R. L., & Lawless, J. L. (2004). Entering the arena? Gender and the decision to run for office. *American Journal of Political Science*, 48(2), 264-280.
- Francia, P. (2003). *The financiers of congressional elections: Investors, ideologues, and intimates*. New York: Columbia University Press.
- Freedman, P., Franz, M., & Goldstein, K. (2004). Campaign advertising and democratic citizenship. *American Journal of Political Science*, 48(4), 723-741.
- Gerber, A. S. (2005). Do phone calls increase voter turnout? An update. *ANNALS of the American Academy of Political and Social Science*, 60(1), 142-154.
- Goff, M. (2007). *The money primary: The new politics of the early presidential nomination process*. Lanham, MD: Rowman & Littlefield.
- Green, D. (2003). Getting out the vote in local elections: Results from six door to door canvassing experiments. *Journal of Politics*, 60, 1083-1096.

- Gronke, P., Koch, J., & Wilson, J. M. (2003). Follow the leader? Presidential approval, presidential support, and representatives' electoral fortunes. *Journal of Politics*, 65(3), 785-808.
- Herrnson, P. S. (1986). Do parties make a difference? The role of party organizations in congressional elections. *Journal of Politics*, 48(3), 589-615.
- Herrnson, P. S. (1988). National party decision making, strategies, and resource distribution in Congressional elections. *Western Political Quarterly*, 42, 301-323.
- Herrnson, P. S. (1992). Campaign professionalism and fund raising in Congressional elections. *Journal of Politics*, 54(3), 859-870.
- Herrnson, P. S., Lay, J. C., & Stokes, A. K. (2003). Women running "as women": Candidate gender, campaign issues, and voter targeting strategies. *Journal of Politics*, 65(1), 244-255.
- Hogan, R. E. (2004). Challenger emergence, incumbent success, and electoral accountability in state legislative elections. *Journal of Politics*, 66(4), 1283-1303.
- Johnson Cartee, K. S., & Copeland, G. A. (1991). *Negative political advertising: Coming of age*. Hillsdale, NJ: Lawrence Erlbaum.
- Kahn, K. F., & Kenney, P. J. (1997). A model of candidate evaluations in Senate elections: The impact of campaign intensity. *Journal of Politics*, 59(4), 1173-1205.
- Kolodny, R., & Dwyre, D. (1998). Party orchestrated activities for legislative party goals. *Party Politics*, 4(3), 275-295.
- Maestas, C. (2003). The incentive to listen: Progressive ambition, resources, and opinion monitoring among state legislators. *Journal of Politics*, 65(2), 439-456.
- Magleby, D., Monson, J. Q., & Patterson, K. D. (Eds.). (2007). *Dancing without partners: How candidates, parties, and interest groups interact in the presidential campaign*. Lanham, MD: Rowman & Littlefield.
- Mayer, W. G., & Busch, A. E. (2004). *The front loading problem in presidential nominations*. Washington, DC: Brookings Institution.
- Norrander, B. (2000). The end game in post reform presidential nominations. *Journal of Politics*, 62(4), 999-1013.
- Ornstein, N., & Mann, T. (2000). *The permanent campaign and its future*. Washington, DC: American Enterprise Institute.
- Petrocik, J. R., Benoit, W. J., & Hansen, G. J. (2003). Issue ownership and presidential campaigning, 1952-2000. *Political Science Quarterly*, 118(4), 599-626.
- Sanbonmatsu, K. (2002). Political parties and the recruitment of women to state legislatures. *Journal of Politics*, 64(3), 791-809.
- Schlesinger, A. (1996). *Political parties and the winning of office*. Ann Arbor: University of Michigan Press.
- Seltzer, R., Newman, J., & Leighton, M. V. (1997). *Sex as a political variable*. Boulder, CO: Lynne Rienner.
- Shea, D. M. (1995). *Transforming democracy: Legislative campaign committees and political parties*. Albany: SUNY Press.
- Sigelman, L., & Buell, E. H., Jr. (2003). You take the high road and I'll take the low road? The interplay of attack strategies and tactics in presidential campaigns. *Journal of Politics*, 65(2), 518-531.
- Sorauf, F. (1994). *Money in Congressional elections*. New Haven, CT: Yale University Press.
- Stone, W. J., Maisel, L. S., & Maestas, C. D. (2004). Quality counts: Extending the strategic politician model of incumbent deterrence. *American Journal of Political Science*, 48(3), 479-495.
- Trent, J., & Friedenberg, R. V. (2008). *Political campaign communication: Theory and practice*. Lanham, MD: Rowman & Littlefield.
- Wattenberg, M. P., & Brians, C. L. (1999). Negative campaign advertising: Demobilizer or mobilizer? *American Political Science Review*, 93(4), 891-899.

93

POLITICAL SOCIALIZATION

VANESSA P. BOUCHÉ

The Ohio State University

Research on political socialization commenced in earnest in the 1950s. In an attempt to understand the decision-making process of the American voter, voting behavior scholars in the 1950s found that factors outside an individual's control influenced, indeed dictated, his or her vote choice in a given election. Survey research subsequently revealed that political orientations and opinions were not hastily made or haphazardly decided; rather, they were the result of a long process that seemed to begin early in childhood, a process called *political socialization*.

Among the first groups of voting behavior scholars is the Columbia school. These scholars surprisingly stumbled on socialization in their research on vote choice. In several localized studies (including Erie County, Ohio, and Elmira County, New York), they followed voters through a campaign to examine the influence of the campaign on the vote (Berelson, Lazarsfeld, & McPhee, 1954; Lazarsfeld, Berelson, & Gaudet, 1948). They found that the campaign had little influence on the vote; rather, voters made their decisions well in advance based largely on party identification. Their party identification was based on primary group associations, the most important being the family through the process of socialization. The primary finding of the Columbia school was that voting decisions were determined by social forces having little to do with electoral politics.

Another group of voting behavior scholars, known as the Michigan school, based their research not on localized

studies (which always raises questions of external validity) but on a national representative survey sample. Their substantive results did not significantly differ; they, too, found that the most important factor influencing voting behavior was party identification, which was transmitted through a socialization process from parents (Campbell, Converse, Miller, & Stokes, 1960).

It was in large part the far-reaching influence of these studies that prompted a generation of political scientists to devote their attention to political socialization. The first problem in any research agenda, however, is defining the terms. What is political socialization? How should it be defined? For this, political scientists turned to social psychologists and sociologists for their definitions of socialization. One prominent social psychologist from the 1950s defined socialization as follows:

the whole process by which an individual, born with behavioral potentialities of enormously wide range, is led to develop actual behavior which is confined within a much narrower range the range of what is customary and acceptable for him according to the standards of his group. (Child, 1954, p. 655)

The first key word of this definition is *process*. Socialization is not something that happens overnight; rather, it is a process that takes place over time and space. The second important term is *behavior*. Socialization is a slow, incremental process that is made manifest over time in observable behavioral outputs.

Sociologists were also intrigued by the question of socialization. Also in the 1950s, Herbert Hyman (1959) argued that socialization is “learning of social patterns corresponding to . . . societal positions as mediated through various agencies of society” (p. 25). While the social psychological definition emphasizes the individual’s behavior vis-à-vis a group’s standard of behavior, the sociological definition emphasizes the individual vis-à-vis society, mediated by various agencies (e.g., groups). A combination of these two definitions most comprehensively identifies the components that, in the aggregate, comprise socialization: It is a process whereby an individual learns from various groups or agencies about the society in which he or she lives, thereby behaving in accordance with the norms and patterns of that society.

A number of political scientists seized on the notion of socialization and sought to understand it specifically in political terms. Gabriel Almond states that political socialization is “the process of induction into the political culture. Its end product is a set of attitudes—cognitions, value standards, feelings—toward the political system, its various roles, and role incumbents” (Almond, 1960b, pp. 27–28). Harry Eckstein (1963) states political socialization is “a process through which values, cognitions, and symbols are learned and ‘internalized,’ through which operative social norms regarding politics are implanted, political roles institutionalized and political consensus created, either effectively or ineffectively” (p. 26). Roberta Sigel (1965) says that “political socialization refers to the learning process by which the political norms and behaviors acceptable to an ongoing political system are transmitted from generation to generation” (p. 1).

Although political scientists over the last 30 years have developed additional definitions of political socialization, they all tend to be combinations or minor adaptations of these foundational definitions. The most fundamental concepts to bear in mind with respect to political socialization are that it is a process that takes place over time whereby political attitudes (including cognitions, affects, and behaviors) are transmitted from primary and secondary groups (which together form a society) and result in certain political orientations (or political attitudes).

Why Is Political Socialization Significant?

As with any research agenda, it is important to ask, “Why should we study this? Why is it important?” There are a number of reasons why the study of political socialization is crucial. First, it expands the scope of political research from the public sphere into the private. It points to the idea that political orientations and attitudes are firmly rooted in identity. Voting behavior scholars have shown time after time that what would otherwise be considered nonpolitical identities actually dictate people’s attitudes toward the political realm. Race, gender, religion, social class, and region, as well as familial and other secondary associations, all

play a significant role in determining how, why, and to what extent an individual participates in and is knowledgeable about politics. The study of politics is intimately personal; the adage “the personal is political” has been shown to be true, and political socialization research also has proven that the political is personal. Thus, the breadth of political research moves into psychology and sociology. Understanding a democratic society requires understanding individual and group relationships and behaviors that give shape to political orientations.

Second, political socialization may impact political stability because of its cross-generational nature. It may also impact political change when exogenous or other factors lead to party or political realignment. The influence of primary and secondary agencies in shaping political attitudes and orientations over time makes the study of political socialization important since it helps explain the stability of democratic political systems (Almond & Coleman, 1960; Dennis, 1968; Easton, 1965). In short, “Political socialization is a stabilizing influence, but not a stagnating one” (Rose, 1964, p. 80). On the other hand, the study of political socialization may help in understanding systemic or nonsystemic political change (Bender, 1967) and why that change occurs.

A third reason to study political socialization is that it is a window into why and how a nation operates. In gaining clarity on the sources of the public’s perceptions about government and their resulting political behavior, we also gain clarity on the operations of government itself (Froman, 1961; Marsh, 1971). Furthermore, the study of political socialization of the elites, and particularly government leaders and elected officials, also assists in understanding political behavior and policy making at the elite level.

What Factors Impact the Political Socialization Process?

Now that a definitional understanding of political socialization has been established, as well as why it matters in the study of political science, the next logical question relates to the inputs into the socialization process. In other words, what factors impact or catalyze the process to move it forward? There are three inputs (outside of primary and secondary groups that exert direct influence, which are discussed in the next section) that play critical roles in the political socialization process, either directly, indirectly, or both: culture, environment, and personality.

The first input is culture. Culture is embedded within the socialization process; it is the foundation on which, or framework around which, an individual is socialized. Kardinger’s Basic Personality Type is one way of understanding culture’s central role in socialization (Kardinger, 1945). Kardinger argues that a parent’s socializing influence is determined by cultural traditions; parents of different cultures socialize their children differently based in their

cultural traditions. He then argues that these early learning experiences have lasting personality effects. The conclusion, therefore, is that similar cultural traditions promote similar learning experiences, which then promote similarities in personality characteristics among those of a specific culture. Applied to the political realm, political culture dictates the ways in which parents politically socialize their children, producing lasting and stable political systems. An individual's culture, therefore, is one important input into how they are socialized. Subcultures within one society are also important socializing agents. Comparing subcultures within the United States (such as those along the lines of race, ethnicity, class, religion, and gender) is one way of understanding how different groups are socialized (Jaros, Hirsch, & Fleron, 1968). For example, the evangelical political culture might be different from that of Jewish political culture, and identifying sub-cultural differences might help explain politically relevant behavioral and attitudinal differences between these groups.

The environment also impacts the political socialization process. On the one hand, the environment may be construed as a catchall that includes any social agent such as media, education, peer groups, or family (Froman, 1961). On the other hand, the environment may mean the political environment within which a person lived, and the political personalities and events that took place during that political era (Searing, Wright, & Rabinowitz, 1976). Thus, environment is an input variable impacting the socialization process that, when disaggregated, encapsulates many individually distinct variables including, but not limited to, any and all primary and secondary associations, the media, contemporary events, and political personalities.

The third variable that has a significant impact on the political socialization process is personality. Although there are a number of definitions of personality in social science literatures, among the most easily operationalizable definitions is that personality is a composite of traits (Guilford, 1959). Of course, personality is not entirely independent of culture and environment and may be impacted, at least in part, by both. Nonetheless, gauging certain individual personality traits is another way to understand the political socialization process, specifically individual political orientations and from where they are derived.

Although it is relatively simple to break down the factors within the socialization process, pinning down the components of each of those factors, as well as the relationships among those factors, is a complex and difficult task. Theoretical development of political socialization has been relatively stagnant over the last 10 years exactly for this reason. When there are innumerable variables that are constitutive, reversely causal, endogenous, and nested, it is particularly challenging to peel them back individually to develop a parsimonious theory that can be empirically tested with easily operationalized variables. This is, perhaps, the fundamental problem in political socialization

research. A discussion of specific agents of socialization, as well as differences among those socialized, might provide a slightly more comprehensible framework for understanding political socialization.

Who Are the Agents of Political Socialization?

As previously discussed, the agents of political socialization were originally bifurcated by the early voting behavior scholars into two categories: primary groups and secondary groups (Berelson et al., 1954; Campbell et al., 1960; Campbell, Gurin, & Miller, 1954; Lazarsfeld et al., 1948). Primary groups are usually defined as those composed of the nuclear or extended family. Secondary groups are everything else, those groups with which an individual has regular and continuing contact that, over time, exert influence on feelings, thoughts, and behavior. Examples of secondary groups include schools, peer groups, occupation, or the media (Bender, 1967).

Within primary groups, parents and families are thought to most significantly influence the political socialization process for a number of reasons. One key reason is that children first establish a sense of identity and belonging from their parents and family. Given the family's central role in providing basic necessities such as food, clothing, and shelter, as well as emotional needs such as love, acceptance, and confidence, meeting these needs in abundance leads to a socially trusting and politically engaged individual, whereas deprivation of these needs leads to lack of social trust (stemming from a lack of familial trust) and resulting political and social apathy (Davies, 1965).

A second reason families are important is that parents pass on certain personality characteristics. Social trust can be established not only through provision of basic necessities but also through a parent's personality predisposition. Parents' levels of interpersonal trust were found to be the best predictor of children's levels of the same, and parents' levels of personal control were the best predictor of personal control in their children (Renshon, 1975). Thus, through provision of basic needs and through learned or inherited personality characteristics, families play a central role in creating social trust, which is a strong mediating variable in the political socialization process.

The political socialization process passed through the parents is mediated not only by social trust, however. Family structure also plays a role. For example, whether a father is present or absent and whether the father is an authority figure impact the socialization process (Davies, 1965). In addition, the intergenerational trends of grandparents to parents to child also impact the process (Beck & Jennings, 1975). Factors such as the total number of grandparents with a certain party identification, whether the parent is the mother or father, and the strength of partisanship all affect the socialization of children. Another structural component with socialization implications is family

composition, including family size, gender of siblings, and birth order. One study showed that, as family size increased, boys' political efficacy decreased; youngest children have the highest level of political efficacy, and oldest children the lowest; and siblings of the same sex politically disagree more than those of the opposite sex. Thus, there are many other mediating factors beyond social trust that come only from family and which are largely out of a child's control (Broh, 1979).

Given the many mediating factors inherent in the parent-to-child socialization process, the question is whether the parents pass down specific opinions on public policy matters or whether the transmission process is more general, relating to overall beliefs. Most studies argue for the latter, but for different reasons. For example, one study concludes that children do not inherit specific political beliefs from parents, but they do inherit certain notions of what is politically acceptable as a result of the generational correspondence between parent and child (Connell, 1972). Another study agrees, arguing that parents transmit basic beliefs rather than specific opinions and attitudes; however, they argue this is the result of personality traits that are either learned or inherited from the parents, or the "psychic organizations" to which a child is exposed (Renshon, 1975). In short, families play a significant role in the political socialization process because of their central role in providing basic needs, passing on certain personality characteristics, and establishing family structures and organizations. Rather than passing down opinions and attitudes on specific political issues, they pass down general and basic beliefs about what is acceptable.

However, parents and families are not the only players in the political socialization process. Secondary groups cannot be forgotten. Indeed, socializing agencies are significantly more complex than the family alone (Jennings & Niemi, 1968a). Three secondary groups that have gained considerable attention are schools, the media, and contemporary events.

There are many factors within the school environment that can have an impact on political socialization, either strengthening or weakening political efficacy and socialization into the political realm. Some of these factors include size of school, quality of education, curriculum, location of school, and school social composition (Litt, 1963). While one study showed that the high school civics curriculum does not have a significant impact on political socialization (Langton & Jennings, 1968), a different study suggested that school composition does matter (Langton, 1967). The socioeconomic homogeneity or heterogeneity of a classroom can have an impact on the politicization of youth. Working-class students in a heterogeneous class environment are more politicized and are more economically conservative, but they are also more ambivalent toward the political system. Schools are also important in that they are a means to participate in extracurricular activities and within peer groups. One study showed that participation in extracurricular activities gave students positive feelings

toward group integration, which led to feelings of social trust and ultimately positive feelings toward politics (Ziblatt, 1965). Social trust is again found to be a strong mediating factor in the political socialization process, and it can be established through the family, peer group activities, or both.

Another secondary group that may strongly influence political socialization is the media. Different media outlets (print or broadcast) and exposure to various content have differential impacts on political socialization (Hyman, 1963). For example, broadcast news is more strongly related to knowledge about current events than print news, and this generally serves a "compensatory function," providing students that otherwise lack knowledge about current events with the information (Garramone & Atkin, 1986). Other studies suggest that the media is the primary source for political knowledge for most children and actually has an indirect effect on children's political attitudes and behaviors, providing a "vicarious experience" through which to identify with certain political orientations (Conway, Wyckoff, Feldbaum, & Ahern, 1981).

Finally, external events also impact the political socialization process. Although political events cannot be considered a secondary group in the way that school, peer groups, and media can, they are part of that general category in that an individual cannot be analyzed in isolation from the time and space in which he or she lives. Therefore, political events that take place during a person's life span are likely to play some role in his or her socialization process, and the political personalities and events therefore impact political orientations and opinions (Searing et al., 1976). Sears and Valentino (1997) argue, "Longstanding predispositions tend to be socialized episodically rather than incrementally" (p. 45), meaning that periodic political events spur socialization in youth, which develops political predispositions that continue later into life. In a later study, Valentino and Sears (1998) found that when there is significant communication about the political events, the political socialization process is most impacted.

Who Are the Recipients of Political Socialization?

Now that it has been established that primary and secondary groups and events are all agents in the political socialization process, the next logical question is, "Whom are these groups socializing?" The easiest answer is everybody, but scholars have examined a number of different groups based on age, political status, sex, education, race, socioeconomic status, and generations. The vast majority of studies in political socialization have centered around the socialization experiences of children, which is implicit in the literature reviewed previously on the agents of the socialization process, most of which focus on parents, families, and schools. However, the socialization of children

matters only insofar as the attitudes and behaviors learned early in life persist into adulthood, when individuals can exercise their political voice in more official and institutionalized ways. Much of the research assumes this, without actually proving it. Marsh (1971), for example, calls into question the assumptions that adult opinions are the result of political socialization in youth and that adult behavior is determined by attitudes learned during childhood socialization. He argues that these two assumptions, on which the political socialization literature rests, need to be empirically tested before moving forward with a research agenda that is youth-centric.

In addition to the socialization experiences of youth, scholars have looked specifically at political elites to determine if those individuals that run for (and win) elected office are subject to different political socialization experiences than the masses. Marsh (1971) argues that more research should focus on the elite political socialization process because, he believes, it is the elites that disproportionately impact the political system. The results of elite studies have been mixed. Research examining political socialization processes of state legislators concluded that political socialization in childhood is very important for state legislators; however, the same study recognized that a number of other factors, including personal disposition and public events, can also impact the socialization process for these elites (Eulau, Buchanan, Ferguson, & Wahlke, 1959). A later study suggested that political socialization in childhood is not important for political leaders, primarily because it has little bearing on how they respond to their official duties while in office (Prewitt, Eulau, & Zisk, 1966).

A third group that has been examined for differential political socialization processes is based in gender. Although a number of studies suggest that boys are more political than girls (e.g., Greenstein, 1961), the important question for the purpose of understanding political socialization is why this is the case. Hyman (1959) found that girls and boys have different patterns of political learning, which manifests in different political behavior for boys and girls. A more recent study found that individual-level differences in socialization patterns between men and women can help explain the aggregate gender gap in partisanship among the electorate (Trevor, 1999). On the other hand, a different study found “no evidence that the roots of the gender differences in adult political involvement lie in the childhood home” (Burns, Schlozman, & Verba, 2001, p. 141), and although there may be a slight socialization advantage for boys in the realm of political socialization through education, girls are more active in clubs and organizations, which helps to compensate.

Another difference in socialization that has been examined is that between the educated and uneducated, which is related to socioeconomic status and race. For example, one study assessing the differential impact of the high school civics curriculum on black and white students found that the curriculum helped increase political participative measures

for black children from less-educated families, depressed performance and participative measures for black students from better educated families, and overall had a greater effect on black students than white students. The researchers argue this is because “information redundancy” is lower for the black students than the white students, so the underlying factor is the education level of the parents and not race (Langton & Jennings, 1968). A different socialization study looking at the impact of socialization on political stability and regime support showed that both black and white students are compliant toward authority but for different reasons. Black children are compliant because they view authority as being powerful, whereas white children are compliant because they view authority as benevolent (Engstrom, 1970). Finally, a study of Appalachian children found that they are considerably less trusting of government than their counterparts in other regions of the country. The theoretical reason is that an authority figure (father) who cannot provide well for the family does not command respect, and this is then projected onto other authority figures, including government (Jaros et al., 1968). Therefore, education, race, and socioeconomic status significantly impact the political socialization process and therefore perceptions of and trust toward authority, which, in turn, impact political behavior.

Finally, generations are socialized differently. Marvin Rintala (1963) defines a political generation as a group of human beings who have undergone the same basic historical experiences during their formative years, meaning from about 17 to 25 years of age. The expectation is that these historical events and the collective understanding of them shape generations differently. For example, Crittenden (1962) hypothesized that the Great Depression and New Deal had a significant impact on the partisan affiliation of that generation, and Almond (1960a) demonstrated changes in public opinion on foreign policy issues as the United States moved into the cold war.

To summarize, family, school, peer groups, the media, and external events can all influence the political socialization process in many different ways and through a number of mediating factors. And they influence the political socialization of different groups and types of people, with potential socializing differences along age, political status, gender, education, socioeconomic, race, and generational lines.

When Does Socialization Occur?

As previously stated, the majority of studies on political socialization focus on the formative, youthful years. Most of the early scholars based the study of children’s socialization on the assumption known as the primacy principle, that what is learned early in life endures throughout life and into adulthood. Because of the strong influence of family and parents, as well as other socializing agents such as schools and social groups, the experiences of childhood

were regarded as significant in understanding the way adults view the political landscape (Dawson & Prewitt, 1969; Easton & Dennis, 1969; Greenstein, 1965; Hess & Torney, 1967; Hyman, 1959).

However, a study in 1973 tested the primacy principle, as well as the structuring principle, which states that “basic orientations acquired during childhood structure the later learning of specific issue beliefs” (Searing, Schwartz, & Lind, 1973). They found little evidence for either the primacy or the structuring principle and therefore argue that adult socialization experiences are really what need to be examined. They postulate that adult socialization either mediates or replaces childhood socialization in the acquisition of issue beliefs.

Clearly, there are elements of both childhood and adult socialization that help an understanding of people’s political beliefs, feelings, and behaviors. That is why others have stressed the fact that socialization can occur at any-time throughout the life cycle. “Socialization, like learning, goes on throughout life. The case of the development of the child is only the most dramatic because he has so far to go” (Parsons, 1951, pp. 207–208). Often times, whether a researcher decides to focus on the formative or mature years depends on the agency. For example, if a researcher is interested in studying the primary agencies (parents and family), then he or she will most likely focus on the formative years; however, if interested in socialization through secondary agencies (e.g., media or professional associations), he or she will look to the mature years (Bender, 1967). Given that socialization can occur at any point during one’s life, the most comprehensive research would consider different age groups moving through different social and political contexts (Cook, 1985).

How Does the Political Socialization Process Work?

There are yet many unanswered questions as to how the process works. Is the process susceptible to abrupt changes based on changes in family structure, changes in life events, or changes in political events and personalities? Are there different socialization processes for learning cognitive versus affective components of political attitudes? Is knowledge and cognition socialized more from secondary agencies, and is affect socialized more from primary agencies? Finally, does the process move forward in a set sequence (e.g., first affect, then behavior, then cognition), or is it random?

Although these are some questions that remain regarding the internal mechanisms of the political socialization process, there have been inquiries into whether the process is latent (unconscious) or manifest (conscious). Bender (1967) argues that latent socialization processes are often those learned via primary agencies in childhood. Youths may not be conscious of the political messages and cues

they receive from family members. Manifest socialization processes are associated with more conscious learning in adulthood from secondary agencies such as work, the media, or other social or professional associations. A study by Almond and Verba (1963) suggests that whether learning is latent or manifest depends on the type of learning taking place. For example, feelings about political authority, in general, are latent rather than manifest, whereas opinions on a given public policy issue are manifest.

What Problems Persist in the Study of Political Socialization?

As with many social science endeavors, there are a number of issues with the study of political socialization that make it a difficult phenomenon to accurately pin down. Five critical issues that complicate the study of political socialization include the following: reverse causality, omitted variable bias, operationalization, standard definitions, and empirical limitations. First is the problem of reverse causality. This means that the political socialization process impacts the political system, but the political system also impacts political socialization. Individuals are not socialized in isolation of the political system; rather, the political system itself is integral in developing political dispositions. The study of political socialization is intended to provide a deeper understanding of system maintenance and stability, but given that political socialization is both an input and an output of the political system, the full extent of socialization cannot be accurately assessed (Dennis, 1968). Therefore, the causal arrow that points from political socialization to system stability points right back at political socialization, making it extremely difficult to disentangle directionality and real effects.

The second problem is omitted variable bias. The political socialization process has a myriad of mediating and moderating variables playing greater or lesser roles at different points in time. The process can depend on whether the socializing agent was a primary or secondary agent, what type of agent, the influence of the agent over the individual, the individual’s personality, cultural factors, and a number of demographic and identity factors, including race, gender, and religion. Given that everything from different types of media exposure to the number of grandparents with the same party identification impact the socialization process, it is likely that there are a good number of variables that are being omitted, thereby biasing the results in favor of the variables that are included. Although omitted variables will always be a problem in social science research because it is impossible to control for everything, it is particularly acute for political socialization research because both nature and nurture play a role, and both nature and nurture constitute innumerable factors and variables.

Third, operationalizing abstract variables is also problematic. Although this is not unique to political

socialization research, the question of how to measure different variables remains an issue. In a model of political socialization, there are inputs (who or what socialized the individual and when), and there are outputs (what was socialized). Two researchers could focus on the socializing effect of parents on children, looking specifically at party identification. Because there are many interpretations of the phrase *the socializing effect of parents on children*, the researchers can measure it however they want. One might choose to interpret it based on how often the family spent quality time together and therefore measure it based on how often the members ate family meals. The other research might choose to interpret it based on the family structure and thereby measure it based on whether the mother worked. These are two very different operationalizations of the socializing effect of parents on children.

This leads to the fourth problem of definitions. Within the political socialization literature, a consensus has yet to be established on clearly defined terms. All definitions of political socialization agree that it is a process, but beyond that there is little agreement on exactly how to define political socialization and all its component parts. Establishing clearly defined terms on which scholars in the field can agree would go a long way in building the framework around which a strong research agenda can be built.

Finally, there are empirical limitations to the study of political socialization. Because political socialization is a process that takes place throughout one's life cycle and is impacted by primary and secondary agents, as well as political events and personalities, a comprehensive study would use longitudinal data, following a group of individuals from youth through adulthood and periodically surveying them to perceive stability or change based on specific influencing factors throughout their lives. The only longitudinal study is the Political Socialization Project, which began in 1965. This project, though still a work in progress, surveyed parents three times, the children (who were youths in 1965) four times (into middle age), and the grandchildren once (Jennings, 2000). This sole longitudinal study is critical to research on political socialization because the vast majority of research has been cross-sectional, which demands significantly fewer resources in terms of time, administration, and money.

Future Directions

The research agenda in political socialization that began in the 1950s and boomed through the 1970s came to a screeching halt in the 1980s and, with few exceptions, has remained relatively stagnant since then. Yet there is still territory to excavate and jewels to be mined. Cook (1985) argued that the dearth of research in political socialization is due to the weak theoretical foundation. He avers that theory should be inclusive of environmental factors, rather than merely individual factors, and that all ages and levels

of development should be considered, as well as all sources of learning, including how individuals receive, understand, and process messages. Thus, future work on political socialization should focus on theory building, and because significant empirical groundwork has been established, theory building can incorporate what is already known. Furthermore, theory building should consider cognitive-biological approaches to political socialization (Peterson, 1983). Given that political socialization is a process of political learning, a focus on cognitive developmental and biological forces most likely play an important part.

In addition to theory building, future research in political socialization should incorporate different empirical approaches, as well. The Political Socialization Project is an important step in this direction. Cross-sectional studies are limited in what they can infer about the political socialization process over time and space. Longitudinal studies are therefore critical. In addition, experimental design has not been used as a methodology to better understand political socialization. There is a lot of promise in experimental methodology as applied to various aspects of the political socialization process. Rosenberg (1985) suggests the use of clinical experiments, moral-judgment interviewers, and free-association interviews to forge new methodological frontiers in the study of political socialization.

In short, the future of political socialization research is wide open for new scholars to develop strong, testable theories and to use diverse methods of testing these theories, many of which have yet to be attempted. It is fertile soil waiting to be tilled.

References and Further Readings

- Almond, G. A. (1960a). *The American people and foreign policy*. New York: Praeger.
- Almond, G. A. (1960b). Introduction: A functional approach to comparative politics. In G. A. Almond & J. S. Coleman (Eds.), *The politics of developing areas* (pp. 3-64). Princeton, NJ: Princeton University Press.
- Almond, G. A., & Coleman, J. S. (Eds.). (1960). *The politics of the developing areas*. Princeton, NJ: Princeton University Press.
- Almond, G. A., & Verba, S. (1963). Political socialization and civic competence. In G. A. Almond & S. Verba (Eds.), *The civic culture* (pp. 266-306). Princeton, NJ: Princeton University Press.
- Beck, P. A., & Jennings, M. K. (1975). Parents as "middleper sons" in political socialization. *Journal of Politics*, 37(1), 831-87.
- Bender, G. J. (1967). Political socialization and political change. *Western Political Quarterly*, 20(2), 390-407.
- Berelson, B. R., Lazarsfeld, P. F., & McPhee, W. N. (1954). *Voting*. Chicago: University of Chicago Press.
- Broh, C. A. (1979). Adler on the influence of siblings in political socialization. *Political Behavior*, 1(2), 175-200.
- Burns, N., Schlozman, K. L., & Verba, S. (2001). *The private roots of public action*. Cambridge, MA: Harvard University Press.

- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. Chicago: University of Chicago Press.
- Campbell, A., Gurin, G., & Miller, W. E. (1954). *The voter decides*. Evanston, IL: Row, Peterson.
- Child, I. L. (1954). Socialization. In G. Lindzey (Ed.), *Handbook of social psychology, Vol. II* (pp. 665-692). Cambridge, MA: Addison Wesley.
- Connell, R. W. (1972). Political socialization in the American family: The evidence re examined. *Public Opinion Quarterly*, 36(3), 323-333.
- Conway, M. M., Wyckoff, M. L., Feldbaum, E., & Ahern, D. (1981). The news media in children's political socialization. *Public Opinion Quarterly*, 45(2), 164-178.
- Cook, T. E. (1985). The bear market in political socialization and the costs of misunderstood psychological theories. *American Political Science Review*, 79(4), 1079-1093.
- Crittenden, J. (1962). Aging and party affiliation. *Public Opinion Quarterly*, 26(4), 648-657.
- Davies, J. C. (1965). The family's role in political socialization. *Annals of the American Academy of Political and Social Science*, 361, 10-19.
- Dawson, R. E., & Prewitt, K. (1969). *Political socialization*. Boston: Little, Brown.
- Dennis, J. (1968). Major problems of political socialization research. *Midwest Journal of Political Science*, 12(1), 85-114.
- Easton, D. (1965). *A framework for political analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Easton, D., & Dennis, J. (1969). *Children in the political system*. New York: McGraw Hill.
- Eckstein, H. (1963). A perspective on comparative politics, past and present. In H. Eckstein & D. Apter (Eds.), *Comparative politics* (pp. 1-30). New York: Free Press.
- Engstrom, R. L. (1970). Race and compliance: Differential political socialization. *Polity*, 3(1), 100-111.
- Eulau, H., Buchanan, W., Ferguson, L., & Wahlke, J. C. (1959). The political socialization of American state legislators. *Midwest Journal of Political Science*, 3(2), 188-206.
- Froman, L. A., Jr. (1961). Personality and political socialization. *Journal of Politics*, 23(2), 341-352.
- Garramone, G. M., & Atkin, C. K. (1986). Mass communication and political socialization: Specifying the effects. *Public Opinion Quarterly*, 50(1), 76-86.
- Greenstein, F. (1961). Sex related political differences in childhood. *Journal of Politics*, 23(2), 353-371.
- Greenstein, F. (1965). *Children and politics*. New Haven, CT: Yale University Press.
- Guilford, J. P. (1959). *Personality*. New York: McGraw Hill.
- Hess, R. D., & Torney, J. V. (1967). *The development of political attitudes in children*. Chicago: Adeline.
- Hyman, H. (1959). *Political socialization*. New York: Free Press.
- Hyman, H. (1963). Mass media and political socialization: The role of patterns of communication. In L. W. Pye (Ed.), *Communications and political development* (pp. 128-148). Princeton, NJ: Princeton University Press.
- Jaros, D., Hirsch, H., & Fleron, F., Jr. (1968). The malevolent leader: Political socialization in an American sub culture. *American Political Science Review*, 62(2), 564-575.
- Jennings, M. K. (2000, October). *Participation as viewed through the lens of the political socialization project*. Paper presented at the conference of the Center for the Study of Democratic Politics, Princeton, NJ.
- Jennings, M. K., & Niemi, R. G. (1968a). Family structure and the transmission of political values. *American Political Science Review*, 62(3), 852-867.
- Jennings, M. K., & Niemi, R. G. (1968b). The transmission of political values from parent to child. *American Political Science Review*, 62(1), 169-184.
- Kardinger, A. (1945). *The psychological frontiers of society*. New York: Columbia University Press.
- Langton, K. P. (1967). Peer group and school and the political socialization process. *American Political Science Review*, 61(3), 751-758.
- Langton, K. P., & Jennings, M. K. (1968). Political socialization and the high school civics curriculum in the United States. *American Political Science Review*, 62(3), 852-867.
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1948). *The people's choice*. New York: Columbia University Press.
- Litt, E. (1963). Civic education, community norms, and political indoctrination. *American Sociological Review*, 28, 69-75.
- Marsh, D. (1971). Political socialization: The implicit assumptions questioned. *British Journal of Political Science*, 1(4), 453-465.
- Parsons, T. (1951). *The social system*. New York: Free Press.
- Peterson, S. A. (1983). Biology and political socialization: A cognitive developmental link? *Political Psychology*, 4(2), 265-288.
- Prewitt, K., Eulau, H., & Zisk, B. H. (1966). Political socialization and political roles. *Public Opinion Quarterly*, 30(4), 569-582.
- Renshon, S. A. (1975). Personality and family dynamics in the political socialization process. *American Journal of Political Science*, 19(1), 63-80.
- Rintala, M. (1963). A generation in politics: A definition. *Review of Politics*, 25, 509-522.
- Rose, R. (1964). *Politics in England*. Boston: Little, Brown.
- Rosenberg, S. W. (1985). Sociology, psychology, and the study of political behavior: The case of the research on political socialization. *Journal of Politics*, 47(2), 715-731.
- Searing, D. D., Schwartz, J. J., & Lind, A. E. (1973). The structuring principle: Political socialization and belief systems. *American Political Science Review*, 67(2), 415-432.
- Searing, D. D., Wright, G., & Rabinowitz, G. (1976). The primacy principle: Attitude change and political socialization. *British Journal of Political Science*, 6(1), 83-113.
- Sears, D. O., & Valentino, N. A. (1997). Politics matters: Political events as catalysts for preadult socialization. *American Political Science Review*, 91(1), 45-65.
- Sigel, R. (1965). Assumptions about the learning of political values. *Annals of the American Academy of Political Science*, 361, 1-9.
- Trevor, M. C. (1999). Political socialization, party identification, and the gender gap. *Public Opinion Quarterly*, 63(1), 62-89.
- Valentino, N. A., & Sears, D. O. (1998). Event driven political communication and the preadult socialization of partisanship. *Political Behavior*, 20(2), 127-154.
- Verba, S. (1964, September). *The comparative study of political socialization*. Paper presented at the American Political Science Association, Chicago.
- Ziblatt, D. (1965). High school extracurricular activities and political socialization. *Annals of the American Academy of Political and Social Science*, 361, 20-31.

94

VOTING BEHAVIOR

C. J. WILLIAMS

University of North Texas

Walter Dean Burnham said, “Electoral politics is not the backdrop; it is the essence, the keystone of the political process. The big issues, such as military, economic, and welfare policy are influenced by the electorate’s opinions” (cited in Neuman, 1986, pp. 1–2). Scholars study voting behavior because it matters for the creation and implementation of public policy in democracies. For the purposes of this chapter, voting behavior can be broken down into two subsections: vote choice and vote decision. Vote choice is defined as forming an opinion in support of one candidate over another. A vote choice must be made before an actual vote is cast. A vote decision is defined as deciding whether to take part in the participatory action of voting. Thus, voting is a two-step process. People must choose which candidates they prefer, and they must decide if they are going to vote at all.

Understanding voting behavior is absolutely necessary today because of the ever-growing number of democracies in the world. Since the 1970s, democracy has been the most widely used form of government throughout the world and has been expanding (see Huntington, 1991). If democracy has become the most widely distributed form of government and the votes of those living in democracies guide public policy, it becomes quite evident why studying voting behavior is important.

This chapter discusses five different theories of voting behavior: (1) the sociological theory of vote choice, which

is based on Berelson and Lazarsfeld’s (Berelson, Lazarsfeld, & McPhee, 1954; Lazarsfeld, Berelson, & Gaudet, 1944) Columbia studies; (2) theories of media and vote choice (Zaller, 1992); (3) rational choice theory, which is an economic theory based argument most notably proposed by Downs (1957); (4) the psychological theory of voting behavior, which is based on the work of Campbell, Converse, Miller, and Stokes (1960); and (5) the theory of social capital, which is a cultural theory most notably posited by Putnam (1993, 1995, 2000).

This chapter ends with a discussion of some recent literature concerning voting behavior as well as some possible avenues for further research. This discussion is of course not exhaustive but designed only to guide the political scientist in training toward a possible area of inquiry.

It is important to understand that this chapter looks only at work on American voting behavior, since it is the most studied nation in terms of voting behavior, although these theories can be applied to vote choices and vote decisions in other countries. Further, this chapter focuses on seminal works that explain voting behavior.

Sociological Theory of Opinion Formation

Berelson and Lazarsfeld (Berelson et al., 1954; Lazarsfeld et al., 1944) propose that “[a person’s] social

conditions and ethnic affiliations, his family tradition, his personal associations, his attitudes on the issues of the day, [and] his membership in formal organizations” (p. 37) determine a person’s vote choice. Using a panel study in Elmira, New York, over a number of months in 1948, Berelson et al. (1954) sought to test the basic theory that social conditions affect a person’s support for particular candidates.

Berelson et al. (1954) find that social conditions are integral to the formation of opinions concerning candidates. Identifying with a particular ethnic, religious, or racial group has an effect on vote choice. Class structure also has a major effect on vote choice. Finally, geographic cleavages also differentiate people in terms of the opinions they hold concerning candidates and issues.

The first finding suggests that the wealthiest grouping of Americans is the most likely to support the Republican Party. Most people would have assumed this outcome; however, more interesting is the finding that there is an age cohort effect. That is to say, individuals who came of age during the Great Depression show a much more intense tendency to make vote choices based on socioeconomic standing. Berelson et al. (1954) attribute this finding to social and political conditions at the time of socialization for the younger age cohorts. The societal conditions during the Great Depression caused younger cohorts to see the political environment differently than those before them.

Berelson et al. (1954) further find that Catholics vote differently than Protestants. This finding is not simply a spurious relationship, according to Berelson et al., but rather a factor intrinsic to Catholicism and Protestantism. Living within a Catholic or Protestant society has an effect on vote choice. Furthermore, those associated more closely with Catholicism (that is, practicing Catholics as opposed to nonpracticing Catholics) are likely to make a different vote choice than Protestants. This same relationship was found for other minority groups at the time. This, according to Berelson et al., occurs because of social cleavages. Social interactions within groups affect the vote choices of the members of those groups.

Berelson et al. (1954) further find that children tend to hold the same opinions concerning candidates as their parents and vote for candidates of the same parties as their parents; however, these patterns are strongly influenced by the social context the children find themselves in at the time of opinion formation and voting. If a child holds a social status similar to that of his or her parents, he or she is likely to vote in similar patterns. Conversely, if a child’s social status is not similar to that of his or her parents, he or she is likely to make different vote choices than his or her parents.

In sum, the work of Berelson et al. (1954) argued that social contexts have the greatest effect on vote choices. A person’s interactions with other members of groups and organizations predispose that person to hold a certain set of

opinions and vote for particular candidates and parties. Being a member of a social class; religious, racial, or ethnic group; or even a close-knit family can steer a person toward one ideology and away from another.

Theories of Media and Opinion Formation

Many researchers agree that the average citizen is politically unsophisticated and ill informed (Converse, 1964; Neuman, 1986; Zaller, 1992). Some political scientists believe that media play an important role in forming opinions that lead to particular vote choices (Iyengar & Kinder, 1987; Zaller).

Zaller (1992) argues that when asked their opinion, people do not necessarily report a true attitude but report a vacillating attitude that is salient at the moment of inquiry. Zaller begins with the assumption that different people have different levels of political awareness. Further, those who are more aware of politics expose themselves to the media at greater levels.

An important point of Zaller’s (1992), however, is that often this “elite discourse” provides conflicting messages. According to Zaller, a person who is more politically aware is more likely to receive these conflicting messages but is also more likely to reject those ideas that do not fit with his or her basic value system. Therefore, those who are highly aware of politics are more likely to hold consistent opinions—meaning that they are more likely to hold opinions that fit together into a coherent system of beliefs at one side of an ideological spectrum.

Zaller (1992) argues that those with lower levels of political awareness receive less information from elite discourse through the media but are also less likely to reject beliefs and opinions that do not fit with a particular belief system. Those with lower levels of political awareness tend to have less consistent opinions.

Zaller (1992) also argues that expressed opinions tend to be unstable. He contends that individuals’ opinions are based on considerations that are most easily accessible to the individual at any particular time. When a person expresses an opinion, according to Zaller, he or she is not expressing a deeply felt opinion but is taking a set of considerations that he or she remembers from elite discourse in the mass media and expressing an opinion based on these considerations. If asked for an opinion at another time, a different set of considerations may be sampled and the opinion may change. This means that people hold opinions that can vacillate. Vote choice is then based on opinions that vacillate based on elite discourses that are sampled at the time of voting.

In sum, Zaller (1992) argues (and with empirical evidence derived from National Election Study data finds) that people look to elites in the media for information concerning politics. Those who are more politically aware receive more information from elites but reject

more information. Those with lower levels of political awareness receive relatively less information from elites in the media and reject less of the information they receive. When asked to express an opinion, people take a sample of the information they possess and base an opinion on this.

Zaller's (1992) theory and findings are very parsimonious. The argument that the media greatly affect opinions and vote choice is supported by earlier research as well.¹ Iyengar and Kinder (1987) find support for the role of mass media through media priming, media framing, and agenda setting.² Neuman (1986) presents an argument that supports Zaller's conclusion that media's effects are based on the sophistication of the audience; however, he comes to the conclusion that media have little effect on opinion formation and vote choice. Neuman's theory is predicated on the idea that mass media is entertainment oriented rather than political-communication oriented. Further, Neuman presents evidence that even when people do encounter political communication, they can remember only about 5% of the information they receive. Neuman, however, does agree with Zaller that the politically aware receive more information from the media than those who are less politically aware, but he believes that this greatly constrains the effect of the media, unlike Zaller, who posits that the media still has a large effect, despite the differences of political sophistication.

The sociological theory and theories of media are similar in that they explain vote choice but not vote decision. The next two sections of this chapter discuss two theories that explain vote choice as well as the causes of vote decision.

Rational Choice, Opinion Formation, and Voting Behavior

Rational choice theory argues that people use a cost-benefit analysis in making decisions. This perspective was first popularized by Downs (1957). Vote choices and vote decision, according to Downs, are rational.

Rationality is an often-misunderstood term. It means that individuals possess ordered goals and they seek those goals in the most efficient fashion possible. In terms of vote choice, those ordered goals lead to a person preferring a particular candidate or policy. In terms of vote decision, a person should vote if he or she gains more utility from voting than from not voting.

When choosing which candidate for whom to vote, individuals place candidates on a continuum that ranges from left to right (liberal to conservative ideology), and after evaluation of the candidate's opinions and stances, the individual actors then aggregate their own preferences and beliefs and place themselves on the same continuum. The individual actors then decide to support the candidate who is closest to them on the continuum.

A notable adjustment to the conventional theory of rational vote choice is the directional theory of issue voting (Rabinowitz & MacDonald, 1989). In this theory, Rabinowitz and MacDonald argue that people still place themselves and candidates on a continuum; however, people do not necessarily choose the candidate that is closest to themselves. Rather, Rabinowitz and MacDonald theorize that individuals hold a particular affinity toward one direction of policy making, which is right or left. The basic idea of Rabinowitz and MacDonald's model is that people act in a rational manner; however, if a person has an affinity to more leftist policy, he or she will not support a rightist candidate, even if that candidate is closer to his or her position on the continuum than the closest leftist candidate. Rabinowitz and MacDonald argue that people will not jump the center point of the continuum and support a candidate on the other side of the continuum.

Understanding how rational choice leads to vote decisions requires a more in-depth look at the theory of rational choice. The rational choice perspective assumes that individuals possess complete information concerning the costs and benefits of voting. The costs of voting are any type of utility expended in preparing to vote or actually voting. The benefits of voting are any types of utility gained from the election of preferred candidates. Deciding to vote, however, is not simply a calculation of the costs and benefits but also the likelihood of making a difference in an election. This is because one individual does not have the final say in who takes office and who does not. Each individual is one of many voters who collectively decide who will lead. Therefore, the likelihood that the individual will make a difference must be taken into account. Thus, the voting calculus first posited by Downs (1957) is simply stated since the probability of an individual voting is a product of the likelihood that one will make an electoral difference times the benefits one will receive from voting minus the costs of voting.

A vote decision requires that a person first know the costs and benefits of voting as well as the likelihood of making a difference. The costs of voting are quite varied. Monetary costs of voting can range from the cost of the gasoline needed to drive to a polling place to taking time off from work to vote. Temporal costs of voting are any situations in which one must take time from one activity to engage in a political activity, such as gathering information on candidates in order to make an informed decision. In addition to the time and energy spent informing oneself about candidates and the actual act of voting, costs include any effects of governmental outputs that decrease an individual's utility. For example, if an individual receives most of his income from Social Security and a candidate is elected who desires Social Security reform that will decrease payments, the individual will lose money, which is a cost.

The benefits of voting are derived from governmental outputs that increase an individual's utility. To estimate the

benefits derived from governmental outputs an individual must estimate what each candidate will do for him or her if elected. A person decides if a candidate will supply him or her with greater utility by placing the candidate on a left–right continuum, as discussed previously. If a candidate is close to the individual on the continuum, the individual's utility to be received is estimated to be high. If the candidate is far from the individual on the left–right continuum, the future utility of the individual is estimated to be low. After estimating the cost and benefits of voting, the potential participant then must also estimate the likelihood of making a difference in an election. If the benefits multiplied by the likelihood of making a difference in an election still outweighs the costs of voting, an individual will decide to vote: a positive vote decision.

Rational choice theory has been heavily criticized, however (see Green & Shapiro, 1994; Stokes, 1963). This theory makes a number of strict assumptions that may or may not be accurate: individuals possess full information, they are able to place candidates on a political spectrum, and they are able to place themselves on a political spectrum. Further, the collective action problem presents a major theoretical hurdle for rational choice theorists.

Collective Action

Olson (1965) posited that people will not engage in an activity when they can receive the same utility regardless of engaging. Olson's theory is important in terms of vote decisions because an individual's one vote most likely does not affect the outcome of an election. Therefore, the outcome does not depend on the individual voter. That is to say people receive roughly the same utility regardless of whether they participate. The collective action problem, thus, is that individuals will not vote if they are rational.

Within the rational choice perspective, this means that no one should vote. However, one sees that people consistently vote in many elections. There must be an explanation of why individuals vote when, rationally speaking, they should not.

The D Term

One of the most widely cited approaches used to explain the participation paradox is Riker and Ordeshook's (1973) addition of the "D term," or duty, into the voting calculus. This explanation supplies a benefit to overcome the costs of voting that is not associated with utility gained from the election of a particular candidate. Any pleasurable experience derived from the action of voting can be labeled the D term as long as it is not a benefit received from the election of a preferred candidate. The D term creates an extra benefit to participation that helps overcome the immense costs of voting and preparing to vote and thus causes people to make a positive vote decision.

Group Membership

Another often-cited explanation for voting despite its inherent irrationality is group mobilization. Group mobilization theory is premised on the idea that people belong to groups and organizations that range from families to work places to unions to churches (Rosenstone & Hansen, 1993). Group mobilization refers to when these groups generate participatory behavior among their members.

Mobilization involves groups pressuring members to participate and specifically participate in support of one particular candidate or issue (Rosenstone & Hansen, 1993). Mobilization creates a situation in which the benefits of voting increase regardless of the benefits one receives from government outputs. By voting, an individual receives acceptance from his or her peers, which is a benefit. Thus, voting becomes a rational decision because social acceptance is a highly desirable benefit, and the costs of voting are outweighed by the benefits.

Informational Shortcuts

As stated previously, informing oneself about candidates is a major temporal cost of voting. If the costs of voting are reduced rather than the benefits of voting increased, voting becomes rational. Reducing the informational costs of participation, therefore, can explain why people vote despite its irrationality. Popkin (1991) argues that various heuristics or shortcuts are used by individuals in evaluating, obtaining, and storing information to decrease the costs of participating. Individuals, according to Popkin, pick up information through everyday interactions and media. Prospective voters can take this small amount of information and apply it in such a way as to form an opinion about each candidate, effectively reducing the costs of participation and explaining why individuals overcome the irrationality of voting.

Procedural Rationality

Besides the collective action problem, another major criticism of Downs's (1957) rationality theory is that individuals rarely have complete information concerning candidates and governmental outputs. McKelvey and Ordeshook (1987) argue that individuals make the best decisions they can, based on imperfect information. Simon (1976) refers to this as procedural rationality.

Procedural rationality makes many of the same assumptions of traditional rationality. Those acting in a procedurally rational manner still seek to increase their utility, still order their preferences, and seek to achieve their ordered goals in the most efficient way available to the actor; however, procedural rationality does not require only one course of action. Procedural rationality means that behavior "is the outcome of appropriate deliberation" (Simon, 1976, p. 131). This means that an individual who

wants to reach Goal A may have more than one path to that goal because, due to uncertainty and incomplete information, the most efficient path may not be the clearest path. As long as a person puts thought into his or her actions rather than simply acting impulsively, the action may be considered rational. Procedural rationality is much like traditional rationality except that it creates more realistic paths for actors, but it also assumes that actors will not engage in an activity that they realize is more costly than beneficial. If behavior is an “impulsive response to affective mechanisms” (Simon, 1976, p. 131), it is considered irrational. Therefore, it is possible to have imperfect information and maintain rationality, as long as the individual has made an effort to make the most efficient decision based on the information available. However, if the individual knows that an action will lose him or her utility and still engages in the action, he or she would be considered to be acting irrationally.

In the case of rational voting, if an individual votes and his or her utility increases from voting, it is rational. If the act of voting loses the individual utility, traditional rationality holds that voting is irrational; procedural rationality holds that voting is irrational if the person expects that he or she will lose utility in voting.

Summary

Rational choice theory is a very parsimonious theory.³ Prospective voters make vote choices simply by placing themselves and candidates onto a left–right political continuum. They decide to support whichever candidate is closest to their own position on the continuum. The prospective voters then make a vote decision based on whether the benefits they receive from voting for the candidate they support outweigh the costs of voting. If the benefits do outweigh the costs, the prospective voters will vote. If the costs are greater than the benefits, prospective voters will not vote.

Psychological Theory of Opinion Formation and Voting

Another theory that explains both vote choice and vote decision is the theory often referred to as a psychological theory. The most notable proponents of this theory are Campbell et al. (1960). They argued that people hold a number of attitudes concerning issues and candidates. They specifically discussed six attitudes: (1) how one feels about the Democrat, (2) how one feels about the Republican, (3) how well each party manages governments, (4) how well each party manages group interests, (5) how well each party manages domestic policy, and (6) how well each party manages foreign policy. These attitudes, which are derived from party identification, guide a person’s vote choice. These different attitudes,

according to Campbell et al., affect which candidate a person supports, as would be expected positive attitudes toward a particular candidate, a particular party, and a candidate’s policies lead one to support that particular candidate.

Campbell et al. (1960) argue that people have psychologically predisposed attitudes, and these are the guiding factors when making vote choices. They find that issues are almost completely inconsequential when people make vote choices. Individuals hold positive and negative attitudes toward the candidates, parties, and other political institutions, which in turn guide their vote choices.

Individuals develop long-lasting psychological connections to political parties (Campbell et al., 1960). In turn, these psychological connections to political parties color people’s attitudes as they are forming. The coloration by the psychological connection to political parties in turn causes people to develop attitudes that are later used to form vote choices. Therefore, the psychological connection to a political party affects a person’s vote choice.

Campbell et al.’s (1960) entire theory hinges on a psychological and stable connection to a particular political party. Rational choice theorists could not explain stable party identification. However, Fiorina (1981) proposed a possible explanation within the rational choice perspective.

Fiorina (1981) argued that people do not forge psychological connections to political parties, but rather, they engage in “retrospective voting.” This means that instead of party identification coloring people’s vote choices, people are evaluating previous party and candidate performance and deciding if a particular party has increased their own utility or decreased it. Voters keep a running tally in their minds of which party has increased their own utility. The connection researchers observe between a voter and a party is a person’s positive tally for that party. This, however, also means that party identification can change as people recalculate their tallies. If this is the case, party identification does not color attitudes and Campbell et al.’s (1960) explanation of vote choice may not be accurate.

Regardless of whether Campbell et al. (1960) or Fiorina (1981) is correct concerning the psychological origin of vote choice, the question still remains: What causes a person to take action on these opinions and vote?

Campbell et al. (1960) find that when attitudes are in conflict, individuals are much less inclined to vote. If people decide for whom to vote based on psychologically based attitudes and these attitudes are in conflict, they will logically have a difficult time making a vote decision.

Other factors were also found to affect an individual’s vote decision. One of Campbell et al.’s (1960) major findings is that the intensity of partisan preference is a major factor in determining turnout. Those who strongly support a party are more likely to vote. Further, those who strongly support a party are more likely to feel an importance in voting. These two findings fit together

very easily. Those who feel a strong level of support for one party are more likely to view voting as important and therefore are more inclined to vote. This is intuitive. People are more likely to engage in an action they deem to be important, whereas if a person does not think something is important, he or she is much less likely to engage in that action. Moreover, it makes sense that people with a strong party identification are more likely to see voting as important. If a person has a psychological connection to a party, he or she would, logically, also want that party to succeed.

Another important finding of Campbell et al. (1960) is that those with a strong party preference are more likely to vote in close elections. When an election is seen as one-sided by the potential participant, the strength of attitudes does not matter in determining whether that potential participant votes. However, if the individual sees an election as close, the intensity of party preference is quite important in determining whether the individual votes.

Campbell et al. (1960), however, admit that partisan strength is not the only psychological factor affecting a person's vote decision. Campbell et al. cite a number of other factors that are highly correlated with voting. The first is interest in a campaign. Campbell et al. find that those who say they are interested in a campaign are also more likely to vote. It is important to note that interest in campaigns does not necessarily lead to voting; it is possible that a person is interested in a campaign because he or she has already decided to vote.

A second additional explanatory factor that Campbell et al. (1960) posit is that caring about the election outcome matters in the vote decision. Simply stated, those who have an electoral preference are more likely to decide to vote.

Campbell et al.'s (1960) third additional explanation concerns political efficacy. If people feel that government is responsive to them, they are more likely to engage in politics and therefore vote. This is a long-term, psychological explanation, rather than short-term interest in a campaign and caring about the electoral outcome. Feelings of political efficacy tend to be stable and, according to Campbell et al., affect a person over many years rather than on an individual electoral basis.

The final explanation Campbell et al. (1960) present is the idea of civic duty. Campbell et al. argue that the norm of political action becomes psychologically ingrained in the individual. Once this happens to a person, he or she instinctively engages in participatory action.

In sum, Campbell et al. (1960) argue that those who have strong partisan preferences are more likely to vote. Further, interest in campaigns, caring about the outcome of an election, feelings of political efficacy, and civic duty all inspire people to vote.

Just as the sociological and media based theories are similar because they explain only vote choice, both the psychological and the rational theories are similar because

they explain both vote choice and vote decisions. The fifth theory discussed in this chapter, social capital, focuses only on vote decisions.

Social Capital and Voting

Putnam's (1993, 1995, 2000) theory of social capital is quite simple. The basic idea of Putnam's (1995) theory is that social interactions develop social capital.⁴ According to Putnam, organizational membership is the most important type of social interaction in developing social capital. Moreover, social capital is necessary for a properly working democracy. High levels of social capital, according to Putnam, can cause people to engage in participatory action, such as voting.

Using General Social Survey data, Putnam (1995) finds that membership in organizations has declined in the United States. This decline in organizational membership, according to Putnam (1995, 2000), is associated with a decline in social capital. Putnam argues that the decrease in social capital is responsible for the decrease in voter turnout in recent U.S. history. Being a member of an organization builds social capital, which instills in people a desire to participate in governance (1993, 1995, 2000).

Although Putnam's (1995, 2000) evidence seems strong, it is important to note that Putnam's work has been criticized (Jackman & Miller, 1998). Some critics argue that the connection between democratic participation and social capital is not clear. Further criticisms of Putnam's work have focused on other theoretic and methodological problems (see Jackman & Miller). Despite the many criticisms of Putnam's work, social capital is still widely studied.

Recent Research

All of the aforementioned works are seminal and have been expanded on greatly through decades of research. This section attempts to provide a better understanding of where the research concerning vote choice and vote decision stands today, although it is far from exhaustive.

Most research today is based in the aforementioned works but has significantly elaborated on them. One of the most researched theories of vote choice is the spatial model associated with rational choice theory. Jessee (2009) tested the main axioms of the spatial voting model and found that "behavior is in close accordance with the fundamental axioms of the basic spatial voting model" (p. 59). Using a measure of voter ideology based on position taking, Jessee tested the spatial voting model for the 2004 U.S. presidential election. He finds that the spatial model is generally correct but is influenced by partisan bias.

Tomz and Houweling (2008) tested the spatial theory of voting using experimental data. Using ideological

placement in terms of health care policy, Tomz and Houweling find that the proximity model of voting usually associated with rational choice is employed more often than not, although it is not used by all people 100% of the time. Further, they find that the directional model of voting first posited by Rabinowitz and MacDonald (1989) is used the least.

There have also been recent advances in the study of vote choice based on the type of election. Feddersen, Gailmard, and Sandroni (2009) theorized that large elections create an atmosphere where people tend not to vote based on material benefits, as per rational choice theory, but instead vote based on who they believe is the morally superior candidate. As the likelihood of being the difference maker in an election decreases, rational choice theory dictates that people will not vote because they will likely receive the same benefits whether they vote or not. However, Feddersen et al. argue that people can receive an expressive benefit, which is a benefit received from expressing oneself. Feddersen et al. find, using an experimental design, that in large elections people tend not to be self-interested in their vote choice, but rather tend to support the morally superior candidate.

In addition to research concerning vote choice, there has been a large amount of research exploring vote decision. Rational choice theory is still a highly studied paradigm concerning vote decision. Recent research has shown through experimental design that the predictions of rational choice theory do in fact occur (Levine & Palfrey, 2007). They find that voter turnout decreases in large elections, voter turnout increases in competitive elections, and those who support underdog candidates tend to turn out at a greater rate than those supporting the favored candidates.

Further, research concerning vote decisions has recently turned to social psychological theories. Using a field experimental design, Gerber, Green, and Larimer (2008) found that social pressure is causally related to vote decisions. This lends support to the claim that pressure to adhere to social norms plays a major role in causing individuals to vote. Nickerson (2008) also uses a field experiment to test the effect of social interactions on vote decision. He finds that when one person in a household is encouraged to vote, the person contacted as well as others in that same household will tend to vote.

Finally, a recent area of vote decision research has arisen based in genetic research. Findings by Fowler, Baker, and Dawes (2008) show that vote decision, and political participation more broadly, can be at least somewhat attributed to a person's genetic structure. Fowler et al. used a quasi-experimental design to test for differences in vote decision between dizygotic (fraternal) and monozygotic (identical) twins. They find that when twins are identical, they are more likely to vote in the same number of elections than when twins are fraternal, thus establishing that a person's vote decision is at least somewhat influenced by his or her genetic structure.

Future Directions

The study of vote behavior has obviously been a major area of research for quite some time (see Lippmann, 1922). However, this does not mean that all questions concerning vote choice and vote decision have been answered. Vote choice and participatory behavior are still highly studied areas of research today.

Future research concerning vote choice and vote decision may be well served to focus on the differences in behavior in different types of elections, such as Feddersen et al. (2009). Some interesting areas of study may be the differences in party vote choice among elections at different levels of government. Further, researchers may want to explore people's participatory decisions in elections at different levels of government.

Another area of study that needs to be more developed is the line of research initiated by Fowler et al. (2008). There is very little research concerning biological causes of voting behavior. With the development of new genetic and biological science technologies, scholars can likely improve on Fowler et al.'s groundbreaking study. Further, researchers may want to apply theories of biology to questions concerning vote choice. For example, researchers may want to ask if genetic variation accounts for differences in the candidates or parties particular people support.

Finally, it is important to note that voting behavior has been studied quantitatively for more than 60 years. It appears that political scientists have virtually exhausted the questions concerning voting behavior that can be answered using simple survey data. Political scientists have established many correlations between variables, but it is now important for researchers to establish causal links. Research going forward must use interesting and novel research designs to establish causality. Obviously, this must be done using experimental and quasi-experimental research designs.

Conclusion

In the previous sections, a number of different theories concerning voting behavior have been discussed. The theorists who posit the sociological theory of vote choice argue that the social conditions in which people come of age, as well as certain social conditions during adulthood, affect the vote choices people make.

The second theory discussed previously concerns the role of the mass media in vote choice. This theory is based on the premise that people have relatively incoherent opinions and attitudes. However, people receive information as elite discourse through the mass media. Those who pay a large amount of attention to the media receive greater information and reject information that is incoherent with previous information. Those who do not pay a large amount

of attention to the media receive less information and tend to reject less information, therefore creating incoherent opinions. Further, people do not hold static opinions, but rather form their opinions based on the information that is most easily accessed.

The rational choice perspective argues that individuals wish to maximize their utility. This desire causes people to have positive opinions of the candidate running for office that they believe will maximize their utility. This theory relies on having the information necessary to place candidates and oneself on a left–right continuum correctly and interpreting the placement of these candidates correctly. In terms of vote decision, this theory argues that prospective voters decide whether the benefits they receive from voting for the candidate they support outweigh the costs of voting. If the benefits do outweigh the costs, the prospective voters will become actual voters. If the costs are greater than the benefits, prospective voters will not vote.

The fourth theory discussed is the psychological approach. This theory posits that individuals have a psychological affinity toward one party or another. This affinity colors the individual's attitudes concerning candidates. The opinions a person holds concerning candidates for office and political issues are derived from the psychological affinity a person holds toward a particular party, also known as party identification. Further, this theory posits that individuals vote for a number of psychological reasons, the most important being the strength of partisan affiliation. If a person is strongly affiliated with a party, he or she is more likely to vote, unless he or she views an election as one-sided. The psychological theory further postulates that interest in campaigns, caring about the outcome of an election, feelings of political efficacy, and civic duty all inspire people to vote.

The final theory discussed previously concerns social capital. This theory argues that people who belong to organizations and groups build what is known as social capital. In turn, this social capital leads people to engage in participatory behavior, such as voting. This theory, however, has been criticized for not providing a causal link between social capital and participatory behavior.

Notes

1. For more information concerning theories of media and politics, see Chapter 82 in this handbook, titled “Media and Politics.”

2. See Iyengar and Kinder (1987) for an explanation of media priming, framing, and agenda setting.

3. For a more in depth understanding of rational choice theory, see Chapter 5, titled “Rationality and Rational Choice.”

4. For an in depth explanation of social capital and civil society, see Chapter 23, titled “Civil Society.”

References and Further Readings

- Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science*, 37, 246–278.
- Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.
- Berelson, B. R., Lazarsfeld, P. F., & McPhee, W. N. (1954). *Voting: A study of opinion formation in a presidential campaign*. Chicago: University of Chicago Press.
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. Chicago: University of Chicago Press.
- Converse, P. E. (1964). The nature of belief systems in mass publics. In D. Apter (Ed.), *Ideology and discontent* (pp. 206–261). New York: Free Press.
- Downs, A. (1957). *An economic theory of democracy*. Boston: Addison Wesley.
- Feddersen, T., Gailmard, S., & Sandroni, A. (2009). Moral bias in large elections: Theory and experimental evidence. *American Political Science Review*, 103, 175–192.
- Fiorina, M. (1981). *Retrospective voting in American national elections*. New Haven, CT: Yale University Press.
- Fowler, J. H., Baker, L. A., & Dawes, C. T. (2008). Genetic variation in political participation. *American Political Science Review*, 102, 223–247.
- Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large scale field experiment. *American Political Science Review*, 102, 33–48.
- Green, D. P., & Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of application in political science*. New Haven, CT: Yale University Press.
- Huntington, S. (1991). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Iyengar, S., & Kinder, D. R. (1987). *News that matters: Television and American opinion*. Chicago: University of Chicago Press.
- Jackman, R. W., & Miller, R. A. (1998). Social capital and politics. *Annual Review of Political Science*, 1, 47–73.
- Jessee, S. A. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review*, 103, 59–81.
- Lazarsfeld, P. F., Berelson, B. R., & Gaudet, H. (1944). *The people's choice: How the voter makes up his mind in a presidential campaign*. New York: Duell, Sloan & Pearce.
- Leighley, J. (1996). Group membership and the mobilization of political participation. *Journal of Politics*, 58, 447–463.
- Levine, D. K., & Palfrey, T. R. (2007). The paradox of voter participation? A laboratory study. *American Political Science Review*, 101, 143–158.
- Lippmann, W. (1922). *Public opinion*. New York: Harcourt Brace.
- Lipset, S. M. (1959). Some social prerequisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53, 69–105.
- McKelvey, R. D., & Ordeshook, P. C. (1987). Elections with limited information: A multidimensional model. *Mathematical Social Science*, 14, 77–99.
- Miller, W. E., & Shanks, M. J. (1996). *The new American voter*. Cambridge, MA: Harvard University Press.

- Mutz, D. C., & Martin, P. S. (2001). Facilitating communication across lines of political difference: The role of mass media. *American Political Science Review*, 95, 97–114.
- Neuman, W. R. (1986). *The paradox of mass politics: Knowledge and opinion in the American electorate*. Cambridge, MA: Harvard University Press.
- Nickerson, D. W. (2008). Is voting contagious? Evidence from two field experiments. *American Political Science Review*, 102, 49–57.
- Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. Chicago: University of Chicago Press.
- Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 6, 65–78.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- Rabinowitz, G., & MacDonald, S. E. (1989). A directional theory of issue voting. *American Political Science Review*, 83, 93–121.
- Riker, W. H., & Ordeshook, P. C. (1973). *An introduction to positive political theory*. Englewood Cliffs, NJ: Prentice Hall.
- Rosenstone, S. J., & Hansen, J. M. (1993). *Mobilization, participation and democracy in America*. New York: Longman.
- Simon, H. (1976). From substantive to procedural rational. *Methods and Appraisals in Economics* (S. J. Latsis, Ed.). Cambridge, UK: Cambridge University Press.
- Stokes, D. E. (1963). Spatial models of party competition. *American Political Science Review*, 57, 368–377.
- Texiera, R. A. (1987). *Why Americans don't vote: Turnout decline in the United States, 1960–1984*. New York: Greenwood Press.
- Tomz, M., & Van Houweling, R. P. (2008). Candidate positioning and voter choice. *American Political Science Review*, 102, 303–318.
- Uhlman, C. J. (1989). Rational turnout: The neglected role of groups. *American Journal of Political Science*, 33, 390–422.
- Wolfinger, R. E., & Rosenstone, S. J. (1980). *Who votes?* New Haven, CT: Yale University Press.
- Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge, UK: Cambridge University Press.

95

AMERICAN FOREIGN POLICY

JEREMY R. BACKSTROM

University of North Texas

American foreign policy has fluctuated throughout the existence of the United States, stemming from the influence of exogenous dynamics and significant watersheds felt throughout the international system as well as endogenous changes and influences within the American government. Noteworthy declarations such as the Monroe Doctrine, international conflicts such as the Spanish-American War, World War II, and the cold war as well as regional conflicts such as the Vietnam War and the Korean War significantly influenced American foreign policy. Currently, the events of September 11, 2001, represent the major exogenous watershed that influenced the foreign policy decision-making of the U.S. government. In addition to the exogenous dynamics that have been decisive in American foreign policy, the endogenous aspects of the U.S. government such as the president, Congress, the bureaucracy, and American public opinion have considerable influence in foreign policy decision making.

This chapter reviews selected seminal literature regarding American foreign policy and its exogenous and endogenous influences. Although exogenous issues are highly significant to the various objectives of American foreign policy, this chapter only briefly highlights the major exogenous watersheds and focuses more substantially on four endogenous dynamics that impact

American foreign policy and foreign policy decision making. The chapter begins with a brief focus of the shift of American foreign policy from isolationism to internationalism and how the Spanish-American War, World War II, the cold war, and the events of September 11, 2001, influenced this shift of American foreign policy. The chapter then concentrates on the endogenous dynamics, beginning with an analysis of the executive branch and its effects over the direction and decision making of American foreign policy as well as the effects of the controversial War Powers Act on the presidential role as a foreign policy decision maker. Second, the focus shifts to the legislative branch and its reactive role regarding American foreign policy as well as the possibility of congresspersons adopting a proactive role to influence the direction of foreign policy decision making concerning particular issues. Third, the chapter analyzes how bureaucratic politics affect American foreign policy by highlighting the three models introduced in Graham Allison's (1971) classic work, *Essence of Decision: Explaining the Cuban Missile Crisis*, regarding the effects of bureaucracy on foreign policy decision making. Finally, this chapter concludes with a concentration on the effects of public opinion on American foreign policy. This section briefly highlights the theories of classical realist and liberal political thinkers concerning the effects of public opinion on foreign policy as well as a focus on the

variations of the effect of public opinion on foreign policy through a historical analysis.

Exogenous Influences: The Evolution of the Role of the United States in International Affairs

After the American Revolution and the establishment of the United States as a nation-state, the United States conducted itself in a fashion dissimilar to the countries in Europe regarding its role in international affairs. Whereas countries such as England, France, and Spain conducted an interventionist international foreign policy, the United States preferred an isolationist route concerning international affairs and focused primarily within its own borders. In 1823, this strategy expanded from the borders of the United States to incorporate the affairs of the entire Western hemisphere. President James Monroe announced a new shift in American foreign policy, namely the Monroe Doctrine, which established a separate sphere of influence for the United States and the Americas versus the European sphere of influence. Furthermore, the Monroe Doctrine stated that the United States would perceive any attempt by the European powers to exert their influence or establish a colonial presence into the Western sphere of influence as an act of aggression. Finally, the Monroe Doctrine stated that the United States would not interfere within European affairs within their sphere of influence and would not intervene within established European colonies within the Americas. Although the United States did not possess the military means to defend a declaration of this magnitude, the British accepted the assertion of U.S. power since they preferred to drive out the Spanish from the Western hemisphere and viewed this proclamation as a method of attaining this goal. More important, the Monroe Doctrine represents a major shift in American foreign policy, which continued to remain isolationist from the international scope but expanded its sphere of influence to all of the Western hemisphere (Kissinger, 1994; Papp, Johnson, & Endicott, 2005).

The Spanish-American War also represented an important stage in American foreign policy since the entry of the United States in the conflict signified a move of the United States becoming an emerging power in international affairs. Although the United States refrained from becoming an international power implicated in the alliance system and continued to adhere to isolationism and unilateralism after the Spanish-American War, the decision to declare war on Spain had two major implications. First, it denoted the initial entry of the United States into world affairs. Although the United States would continue to internally debate and argue over whether American foreign policy should remain isolationist or

shift toward an internationalist foreign policy for many years, the United States entered into several international conflicts and diplomatic interactions with other countries after the Spanish-American War. Second, in the aftermath of the Spanish-American War, the United States seized possession over the former Spanish colonies of Guam, the Philippines, and Puerto Rico. Formerly, the United States was exclusively an isolationist country; however, now the United States became a colonial power and expanded its interests to other hemispheres (Papp et al., 2005).

Although the United States gained colonial territories in other hemispheres, most Americans did not agree with the idea of U.S. involvement in world affairs. This debate continued during World War I as well as during the interwar years as demonstrated in the aftermath of World War I. President Woodrow Wilson was one of the key figures in the formulation of the Treaty of Versailles and the establishment of the League of Nations; however, Americans, particularly the senators who voted against the ratification of the Treaty of Versailles and the League of Nations, still believed that the United States should adhere to a foreign policy devoted to isolationism and unilateralism whereas the Treaty of Versailles strongly espoused internationalism and collective security (Kissinger, 1994; Papp et al., 2005). The debate over the role of the United States in world affairs persisted into World War II until December 7, 1941, when Japan bombed Pearl Harbor. At this point, the United States entered into World War II, and American foreign policy would never be the same.

With the entry of the United States into World War II, the American foreign policy shifted to one of multilateralism and internationalist in scope. This particularly was demonstrated with the establishment of the United Nations and the passage of the Bretton Woods Agreement and the General Agreement on Tariffs and Trade. Furthermore, the United States signed the North Atlantic Treaty with several other countries, creating the North Atlantic Treaty Organization, a military alliance, on April 4, 1949. The involvement of the United States in the establishment of these organizations and future participation within the organizations signified that American foreign policy had surpassed isolationism and was now firmly entrenched in internationalism and multilateralism (Ambrose & Brinkley, 1997).

After the end of World War II, world affairs had changed immensely as the United States and the Soviet Union emerged as the world powers or superpowers in the international system. Although the countries had been tenuous allies during World War II, competing political ideological and national interest differences soon forced a wedge between the countries, and the international system was greatly affected by the bipolar world established during the cold war. During the cold war, American foreign policy continued to be focused on internationalism and multilateralism, but confronted with an enduring power struggle with the Soviet Union, American foreign policy focused on the policy of

containment as well. Devised by George F. Kennan, the policy of containment sought to thwart the spread of Communism to non-Communist countries throughout the world. As the Soviet Union expanded its influence to satellite countries with the spread of Communism, the United States also increased its area of influence to various countries throughout the world in an effort to counter the Soviets (Ambrose & Brinkley, 1997; Kennan, 1984). By contributing military and economic support to so-called anti-Communist countries, the United States focused on preventing the spread of Communism to other countries throughout the world and averting a domino effect occurring where the loss of one country to Communism would lead to the further loss of neighboring countries to Communism. Although the United States and the Soviet Union did not engage in direct military warfare, the superpowers engaged indirectly through proxy wars in their support of satellite countries. In addition to the indirect interaction through proxy wars, nuclear deterrence became a significant aspect of the cold war as the United States and the Soviet Union accumulated substantial stockpiles of nuclear weapons as a method of ensuring their respective state security. Although both countries possessed the weapons, neither of the countries would employ the use of nuclear weapons because leaders understood the drastic effects to both countries if the weapons were deployed (Ambrose & Brinkley, 1997; Kissinger, 1994; Mingst, 2008).

Although American foreign policy changed to cope with new challenges arising from a post-cold war international environment, the next significant watershed in American foreign policy stemmed from the events of September 11, 2001. After the terrorist attack against the United States by members of al Qaeda under the direction of Osama bin Laden, President George W. Bush responded by shifting American foreign policy to a global war on terrorism, which commenced with the launching of an attack against Afghanistan for harboring bin Laden and al Qaeda. In 2003, the war on terrorism expanded to include the Republic of Iraq since the United States was convinced that Saddam Hussein presented a terrorist threat through the alleged possession of weapons of mass destruction and support of terrorist organizations (Mingst, 2008). Although Hussein was overthrown from power and no weapons of mass destruction were found, American foreign policy continues to be greatly affected by the war on terrorism since the conflicts in Afghanistan and Iraq continue into the second decade of the 21st century.

Endogenous Influences

The President and American Foreign Policy

The U.S. president and executive branch arguably wield the largest amount of power in regard to American foreign policy. Although Congress and the bureaucracy of the national government can certainly affect how foreign

policy is conducted, the executive branch by far has the most control over foreign affairs (Peterson, 1994). Originally, this was not what the framers intended when designing the U.S. Constitution; however, the presidential powers regarding foreign policy have increased significantly since the founding of the United States. Specifically, the framers went to great lengths to ensure that Congress controlled the decision to go to war by conferring the power to declare war solely to Congress. On the other hand, the president exercises widespread authority in times of crisis, but the overall decision to initiate war was left in control of Congress (Edwards & Wayne, 2006; Pika & Maltese, 2008). Therefore, the U.S. Constitution established a shared power regarding issues of war between the executive and legislative branches. Although this shared power was established in the U.S. Constitution, historical precedent demonstrates that the power of the president has been advanced significantly, specifically during times of war and crisis, at the expense of congressional power. Furthermore, after the declaration of war has been made, the president, acting in his role as the commander in chief, is granted extensive power based on Article II of the U.S. Constitution and congressional delegations of authority (Pika & Maltese).

In addition to issues concerning war and crisis, the president has several other powers in regard to foreign policy, such as diplomatic treaties, presidential appointments, and executive agreements. According to the U.S. Constitution, the president may enter into international diplomatic treaties with other countries; however, the president must consult with the Senate concerning the treaty. In addition, the treaty must then be ratified by a two-thirds majority of the Senate. In general, international treaties have been approved without amendment by the Senate, with a few exceptions such as the Versailles Treaty after World War I. This may also be related to the fact that many have been withdrawn by presidents who anticipated defeat due to a lack of congressional support as displayed with President George W. Bush and the Kyoto Protocol.

On March 28, 2001, President Bush declared that the United States would not implement the Kyoto Protocol regarding combating global warming through a reduction of greenhouse gases. Although it appeared that Bush made a dramatic turn in U.S. policy decision making, Congress had not supported this direction for some time. This particularly began in July 1997, when the U.S. Senate unanimously approved Senate Resolution 98, stating that it would not sanction a global climate treaty that would damage the U.S. economy or that failed to compel the reduction of emissions within the same time period for both developing countries and developed countries. Furthermore, even when President Clinton agreed to the Kyoto Protocol, he did not present the treaty to the Senate for ratification because he was aware of the lack of support for the treaty (Fletcher, 2000).

The president may also appoint several key positions subject to Senate confirmation. According to Article II,

Section 2 of the U.S. Constitution, the Senate must confirm presidential diplomatic appointments such as ambassadors to foreign countries but more important the secretary of state, secretary of defense, and the director of central intelligence. This is significant to American foreign policy given that these three positions are highly influential to the course of the administration's foreign policy (Edwards & Wayne, 2006; Pika & Maltese, 2008).

Presidents may also affect foreign policy through executive agreements, which allow them to forego the process of the congressional ratification process. Specifically, an executive agreement is a pact between the U.S. president and the head of state of the other country, which does not necessitate the ratification of the U.S. Senate. Since the presidency of Ronald Reagan, the use of executive agreements has increased immensely, which may reflect an attempt to evade the congressional ratification process (Edwards & Wayne, 2006).

In 1973, Congress attempted to strike back at the rampant conduct of the Lyndon B. Johnson administration during the Vietnam War through the establishment of the War Powers Act. Although the passage of the War Powers Act was delayed until the term of President Richard Nixon, it symbolized an attempt of Congress to strengthen its powers in the realm of foreign policy decision making as well as to establish an effective restraint against the executive branch and its unbridled control in decision making regarding the American foreign policy and the deployment of American troops in a hostile environment. After its passage, President Nixon vetoed the War Powers Act; however, Congress was able to obtain the necessary two-thirds majority to override the veto. Although the necessary majority was achieved, the bill itself was weakened compared with its original overall objectives in the process of attaining a majority in the House of Representatives to override the veto (Fisher, 2006).

Specifically, according to the resolution, the president may deploy American military troops overseas for a 60-day period during peacetime prior to obtaining congressional approval for the action and may appeal to Congress for an extension period of up to 90 days. After the 60-day period, if Congress does not give approval for the deployment, the president has 30 days to extract the troops. It is debatable whether the War Powers Act indeed fulfills its original intentions. Louis Fisher (2006) argues that the 60-day period itself is a much larger expansion of power than the original framers of the bill intended to grant the president. In addition, the resolution requires the president at all possible times to confer with Congress regarding the action prior deployment of the American military into a hostile environment as well as to submit a report to Congress within 48 hours of the deployment. The 60-day window actually begins when the president reports to Congress concerning the deployment; however, it is typically reported in a general manner. This allows for the president to conduct foreign policy endeavors without the advice and consent of Congress, or what Fisher calls a "collective judgment of

the branches" (pp. 279–280). Therefore, it is highly debatable whether the War Powers Act provides a restraint against presidential adventurism and actually satisfies the intentions of the original framers of the resolution. Furthermore, since the passage of the War Powers Act, previous presidents asserted that it unconstitutionally limits them from performing their duties as commander in chief to provide defense. To overcome this limitation, they have interpreted the Constitution in a flexible manner, specifically in regard to the requirements for reporting and consulting with Congress (Edwards & Wayne, 2006).

Congress and American Foreign Policy

Deemed the face of the nation for reasons of international visibility, it is commonly assumed that the majority of foreign policy making is conducted by the president and the executive branch; however, this view overlooks the significant influence and power that the legislative branch wields over American foreign policy. First, as stated in Article I of the U.S. Constitution, Congress has the sole authority to declare war. Second, Congress also possesses the power of the purse, which may be used as a tool to influence the executive branch on how foreign policy is conducted. Third, according to Article II of the U.S. Constitution, the president may negotiate a diplomatic treaty only with the support of a two-thirds majority of the Senate. Fourth, Congress also advises and consents to presidential appointments such as ambassadors and consuls. Through these listed means as well as others, Congress has a significant influence over the direction of foreign policy decision making.

Although in each of these previously listed manners Congress acts in a reactive manner to the foreign policy endeavors of the president and the executive branch, Ralph Carter and James Scott (2009) suggest a theory of congressional foreign policy entrepreneurship in which congresspersons may be proactive in their foreign policy influence ventures; specifically, they may act as what are termed *foreign policy entrepreneurs*. In this manner, the congressperson chooses to lead the development of foreign policy by attempting to influence the foreign policy of a particular issue outside the desires of the current administration for a continuous period of time. Specifically, this means that the congressperson must engage in more than one attempt of foreign policy entrepreneurship.

A congressperson becomes a foreign policy entrepreneur when he or she engages in developing a new policy regarding a foreign policy issue. This decision may specifically occur during a policy vacuum, a period where there is no policy regarding the issue, or during a policy correction, when the congressperson feels that the current policy is flawed. Furthermore, in a proactive approach, the congressperson does not wait for the president to take action regarding a foreign policy matter or inform the party or country on his or her stance regarding the affair; rather, the

congressperson acts on the matter proactively in an attempt to influence foreign policy decision making. Specifically, the congressperson can introduce new legislation regarding the foreign policy issue, offer amendments to existing legislation, conduct policy research, travel to determine the realities of the issue, or hold hearings to publicize the foreign policy issue that he or she is promoting (Carter & Scott, 2009).

Carter and Scott (2009) identify that a congressperson may decide to progress in this fashion for several reasons. First, this pursuit may not be completely influenced by a desire for reelection because foreign policy rarely directly affects the constituency of a congressperson. On the other hand, if a congressperson has a rather large presence of a particular ethnic group within his or her district, this may influence the decision to proceed with actions to influence the foreign policy regarding the respective country of the ethnicity of the constituency. Second, the congressperson may pursue this issue in order to gain respect and influence within Congress. He or she may gain a reputation for expertise in the subject, which may assist him or her in garnering influence in Congress. Third, the congressperson may have a personal policy position regarding the issue. This may stem from core values, personal experiences, and family experiences. The core values of a congressperson may come from his or her respective morality as well as the influence of particular issues that are important to the congressperson. The personal experiences of a congressperson may influence him or her to pursue a particular policy since he or she may have an expertise regarding the issue. Family experiences may also motivate a congressperson to pursue a policy since many have first- or second-generation family members.

To understand the theory of foreign policy entrepreneurship, Carter and Scott (2009) draw on John Kingdon (1995) and his ideas regarding foreign policy entrepreneurs in his work, *Agendas, Alternatives, and Public Policy*. According to Kingdon, a congressperson's decision to engage in foreign policy entrepreneurship may shape policy windows. First, policy making is conducted in the middle of a "problem stream" where those within the government as well as the public recognize a particular problem. During a *problem stream*, the foreign policy entrepreneur must specifically define the problem in order to enhance the awareness of the policy problem to others. The term *policy stream* refers to potential solutions to the policy problem, which are developed by policy experts. Finally, the term *political stream* refers to a change in conditions or specifically to when key actors within the governments and institutions as well as society change in their disposition toward the issue. If the problem stream, policy stream, and political stream correspond at the same time, the policy window will open and the foreign policy entrepreneur will have the opportunity to construct changes regarding the particular policy (Kingdon, as cited in Carter & Scott).

When this process occurs and foreign policy entrepreneurs have an opportunity to create policy changes, the success

of their endeavors is influenced by the structural features to include which house of Congress they serve in, if they hold a position on a policy-appropriate committee, which political party they are members of, whether they hold party leadership positions, and whether they are senior congresspersons. Once congresspersons decide to engage in foreign policy entrepreneurship, they will attempt to frame the policy issue to others in order to gain their support and ensure the success of their endeavor. Furthermore, they will make contact with a senior administration official to seek a change in the particular policy, to determine whether to go public with the issue, or possibly to conduct formal measures in order to seek to adopt legislation regarding the policy matter (Carter & Scott, 2009).

Through the theory of foreign policy entrepreneurs, Carter and Scott demonstrate how Congress can project influence on foreign policy creation in a proactive manner as opposed to a reactive manner as prescribed in the U.S. Constitution and by established precedents, where the executive branch essentially produces the foreign policy initiative. Although the constitutional powers established for the legislature are powerful and influential in American foreign policy, the process of foreign policy entrepreneurship allows the congressperson to endorse foreign policy issues and initiatives that are significant to him or her. Overall, foreign policy entrepreneurship is becoming a more common practice among congresspersons, which will undoubtedly affect the process of foreign policy creation in the future.

The Bureaucracy and American Foreign Policy

In *Essence of Decision: Explaining the Cuban Missile Crisis*, Graham Allison (1971) analyzes the Cuban missile crisis in October 1962 and the consequent naval blockade of Cuba by developing three models or frames of reference in order to highlight the understanding of foreign policy decision making: the rational actor model (RAM), the organizational behavior model (OBM), and the governmental politics model (GPM). Although there is some overlap within the three models, each model represents a distinctive manner of the effect of the bureaucracy on foreign policy decision-making.

The first model is the RAM, where a researcher observes the situation as the state itself as a single entity thinking and acting in unison. In this manner, the state is the key actor and acts in a rational manner. Security is the highest priority of the state and influences various other objectives respective to the state. The state will then select the preference that coincides with the highest of its respective objectives. Therefore, the state is value maximizing in its actions. Allison states that this approach to understanding foreign policy decision making is the most common (Allison, 1969, 1971; Allison & Zelikow, 1999).

Furthermore, the state's decision of proceeding with an action is determined by several factors: objectives, perceptions of other options, an assessment of the consequences

of their actions, and an overall evaluation of each consequence. The culmination of these noted elements will amount to an increase or a decrease to the cost of an alternative option, which ultimately affects whether the decision maker uses the respective option, essentially a cost-benefit analysis. In the end, the decision of the state is viewed as a unified national government action where the state is acting as a single, rational actor in pursuit of the state's national interest. In the RAM, the elaborate interworkings of government and various bureaucratic missions and goals are not stressed on since the state is viewed as the sole significant entity functioning to pursue a cohesive, uncontested national interest (Allison, 1969, 1971; Allison & Zelikow, 1999).

The second model is the OBM. Here, in contrast to the RAM, decisions are depicted as the product of the collaboration of large governmental organizations, which each tend to conduct themselves according to their own standard operating procedures (or SOPs), yielding standard behaviors. Within this model, the single-state decision maker is no longer the key decision-making actor; rather, the loosely associated governmental organizations themselves are the central actor(s) guided by permanent SOPs that are predetermined (Allison, 1969, 1971; Allison & Zelikow, 1999).

In the OBM, Allison highlights the effect of governmental organizations on foreign policy decision making whereby responsibility for various areas of government is divided among the large organizations. Thus, whereas the RAM depicts decision making conducted by the state as a single entity, the OBM displays decision making as an output of numerous relevant governmental organizations, which may be coordinated by government leaders. To organize this process, the governmental organizations developed SOPs, which are previously established for each organization. The organization, however, may evolve over time because of learning or radically change as a reaction to a major crisis facing the country. In other words, rather than a rational decision-making process, according to this model decisions are the consequence of various organizations within the government acting according to their SOPs, or slight modifications of these, in response to problems. Decisions are not depicted as rational, and therefore it is not assumed to be any effort to reach an optimal decision (as in the RAM), but instead decisions are determined by what an organization deems feasible and yet reasonably responsive to the problem (Allison, 1969, 1971; Allison & Zelikow, 1999).

The third model is the GPM, in which decisions are viewed as the outcome of the bargaining among the actors within government. In this model, the leaders and other individuals within the various governmental organizations are actors within the game of bureaucratic politics who seek to advance their objectives, which are formed by national, organizational, and personal goals (Allison, 1969, 1971). Such goals may include service to the nation but also personal objectives such as career advancement. The GPM

depicts politics as a game in which individuals seek to steal the limelight from colleagues at other departments, where subordinates seek to either support or subvert the careers of their bosses, and so on. This model, then, suggests that political decisions are the outcome of complex bargaining games within and across the various organizations that together constitute the national government.

In other words, whereas the RAM viewed decision making as performed by the state and the OBM perceived decision making as the outputs of various governmental organizations, the GPM identifies decision making as a product of bureaucratic politics with the top central organizational leaders as the key players of the game. Within the game, the leaders undoubtedly engage in disagreement, compromise, and bargaining over what direction the government should proceed in regarding foreign policy decisions since each leader has his or her own personal ideas and goals as well as their respective organizational objectives to endorse. Furthermore, the personalities of the leaders become significant to the successfulness of the actor to project his or her position and compel the other actors to agree with his or her position. Allison also suggests that the actor's ability to be successful in the game depends on his or her power, which is a combination of bargaining advantages, the skill and will of the actor in using bargaining advantages, and the opinions of other actors regarding the bargaining advantages and the skill and will of the actor to use them. In sum, as one moves from the RAM via the OBM to the GPM, the image of how decisions are made becomes increasingly more messy and less an orderly and rational process.

Allison's seminal work has spawned subsequent generations of scholars who have produced various theories about foreign policy decision making (e.g., Garrison, 1999, 2001; George, 1980; George & George, 1998; Hart, Stern, & Sundelius, 1997; Janis, 1983; Kowert, 2002). What these theories have in common is a recognition that decision making rarely proceeds as orderly as the RAM would suggest. These theories and frameworks share in common an attempt to better understand how the perceptions and cognitions of decision makers, as well as the dynamics among small groups, among the various branches of government, and interactions with the wider domestic audience shape the decision-making process. Some of these theories are discussed in greater detail in Chapter 40, titled "Foreign Policy Analysis."

Public Opinion and American Foreign Policy

A discussion of American foreign policy cannot be entirely fulfilled without concentrating on the effect of public opinion on American foreign policy. Holsti (1992, 1996) identifies three watersheds that ultimately shaped the relationship between public opinion and American foreign policy: World War I, World War II, and the Vietnam War. The focus on the effect of public opinion truly commenced after World War I, when public opinion played a

role in the decision making of how to create international order in the post–World War I period. The significance of World War II came as scientific polling of public opinion began, allowing for a much more advanced empirical investigation into the opinions of the American public. Finally, the Vietnam War as well as the outcomes after the Vietnam War initiated another focus on the public opinion and its effect on American foreign policy.

Public opinion has been viewed in a disparate manner by liberal and realist classical theorists. Immanuel Kant, Jeremy Bentham, and other theorists suggested that given the structure of a democracy, public opinion provides a safeguard on the foreign policy desires of a government. In this manner, the mystery of foreign affairs is revealed, governments are held accountable for their actions, and public opinion is viewed as the solution to some of the dilemmas of government. However, in a nondemocracy such as a monarchy or a totalitarian government, foreign policy may be conducted without regard to public opinion and in the direction to the whims of the monarch or tyrant. In this manner, there is a lack of constraints on the government, and no accountability is given to the public, allowing the leader to conduct foreign policy in a manner to his or her choosing (Holsti, 1992, 1996).

Conversely, realists perceived public opinion as a challenge to the foreign policy decision making of governments. Whereas the liberal school of thought viewed public opinion as an integral aspect of foreign policy, realists such as Hans Morgenthau and others questioned the ability of the public to contribute to foreign policy decision making. First, they considered the public as being too focused on their daily lives and too far removed from the issues concerning foreign policy. Essentially, the public could not understand the essential issues of foreign policy and consequently could not make an effectual contribution to foreign policy decision making. Second, the public was viewed as irrational, passionate, and emotional, which would lead to ineffective decision making and could possibly jeopardize the country if allowed to contribute to foreign policy. Finally, realists viewed the actual process of foreign policy decision making and diplomacy as one based on secrecy, accommodation, and speed. To have the public involved in the process of foreign policy decision making would be counter to the listed necessary traits and could endanger the state itself or the international system within which it resides (Holsti, 1992, 1996).

After World War I, President Woodrow Wilson hoped to develop a new world order centering on democracy and diplomacy among countries. In this vision, Wilson and his cabinet possessed a liberal outlook on the relationship between public opinion and American foreign policy. In holding this viewpoint, they anticipated a significant and increased role for the public in diplomacy and foreign affairs. On the other hand, other figures such as Walter Lippmann, a journalist, held an opposing view of public opinion and its relationship with foreign policy than Wilson and his cabinet. Espousing realist rhetoric, Lippmann

wrote in his critiques of liberalism that the public was too interested in their daily lives and fulfilling their most pressing needs and desires to become informed about foreign policy affairs. Furthermore, the public was too far removed from the events of foreign affairs to develop an informed opinion of the issue, and the media was unable to provide for this gap of knowledge (Holsti, 1992, 1996).

The period of World War II and its aftermath produced an intensification of scientific public-opinion polling. During this period, polling organizations attempted to ascertain the opinion of the American public regarding a major concern of U.S. policymakers: Should the United States remain an isolationist from world affairs, or should it become an active player? From the information attained from public-opinion polling during the post–World War II period and prior to the Vietnam War, three main ideas were suggested. First, public opinion is highly volatile. In their writings, Lippmann (1925) and Gabriel Almond (1950) depicted the public as passion driven, moody, and volatile. In addition to this observation, polling data illustrated an American public that was surprisingly ignorant of facts and information essential to foreign affairs. Second, public opinion lacks structure and coherence. Early research and studies predicted that the American public would fluctuate between support for the internationalist viewpoint and support for the isolationist viewpoint regarding foreign affairs, similar to the liberal viewpoint versus the conservative viewpoint regarding domestic affairs. In 1964, however, Philip Converse published a study that suggested a different finding. Contrary to the previous belief that a political spectrum of foreign policy support existed, Converse stated that the American public lacked a coherent structure in their political beliefs and typically, their beliefs held a brief impact on their views regarding foreign affairs. Third, public opinion has limited impact on foreign policy. In the immediate post–World War II period, policymakers were split on the liberal–realist divide concerning the effectiveness of the American public opinion on foreign policy. However, in the 1960s, policymakers now viewed public opinion as having little to no impact on foreign policy decision making. In fact, studies during this period displayed that the opinion of constituencies regarding foreign affairs had little influence over their congresspersons, and other studies proposed that the president had an unbridled impact on foreign policy decision making (Holsti, 1992, 1996).

The Vietnam War provided the third watershed regarding public opinion and American foreign policy. Within this period, the relationship between public opinion and American foreign policy gained a renewed significance as many realists including Lippmann questioned the belief of an imperial presidency and a limited impact of public opinion on foreign policy. Furthermore, more polling organizations with a narrower focus in contrast to the general surveys of Gallup were created and devised their surveys in a simpler yet more extensive and in depth manner. Specifically, these surveys focused on public opinion regarding foreign policy in Vietnam, which revealed support for the administration's

foreign policy endeavors yet also supported an end to the Vietnam War (Holsti, 1992, 1996).

Through the findings of the new public opinion information, the three main ideas criticizing the American public were challenged. First, regarding the idea that the American public opinion is volatile, new studies found that the preceding research and surveys were conducted in a flawed manner that may have been the source of the earlier propositions. By modifying their methodology and research approaches, the new studies found that the American public was remarkably stable in their opinions regarding foreign policy yet remained poorly informed concerning facts, geography, politics, and foreign policy. Second, the claim that the American public lacks structure and coherence also suffered from similar methodological issues, and once researchers modified their methodological approaches, this created a split in support for and challenges against Philip Converse's (1964) work on mass belief systems. Furthermore, several studies have questioned Converse's results and found that public opinion does not adhere to an internationalist-isolationist dimension. Since this point, numerous studies have suggested various types of foreign policy attitudes rather than the earlier internationalist-isolationist dimension. Current research has also found that although the American public typically lacks complete information concerning foreign policy, they use simple heuristics in order to compensate for the incomplete knowledge (Holsti, 1992, 1996).

Regarding the claim that public opinion has limited impact on foreign policy, when evidence arose of a relationship between public opinion and American foreign policy, many scholars and analysts continued to disagree that there was a connection between the two. They continued to remain steadfast to the argument that public opinion has no impact on foreign policy, and if any relationship existed, it could be explained as an attempt by the executive to direct public opinion in support of their foreign policy aspirations. To an extent, this was true; however, there were numerous cases of unsuccessful bids for public support. Furthermore, this did not explain the converse relationship. Several studies suggested that in the midst of foreign policy decision making, presidents often considered the postaction response of the American public, which influenced their decisions. This has also been a factor in foreign policy decision making since the public may vote retroactively, and a foreign policy failure may doom the incumbent candidate's bid for reelection. In addition, according to many public opinion officials for the U.S. government, the public has not been viewed as an entity that may be influenced; rather, public opinion has been a significant explanatory variable in presidential decision making regarding foreign policy. Although the causal linkage between public opinion and American foreign policy has yet to be conclusively established, it has been demonstrated that public opinion has an impact on foreign policy decision making (Holsti, 1992, 1996).

Robert Entman (2004) also advocates a noteworthy relationship between public opinion and American foreign policy. Where earlier models suggested that either public opinion has no effect on American foreign policy or a national government holds influence over public opinion and Holsti (1992, 1996) suggested that the American public has an influence over foreign policy decision making, Entman proposes a dual relationship between the concepts. Specifically, he implies that there is a simultaneous interaction between the concepts since the president attempts to sway the public to agree with his foreign policy endeavors while the public also interacts with the president as they inform the executive branch what they are prepared to tolerate regarding the American foreign policy endeavors. In this model, the interaction between the national government and the American public is not a top-down or bottom-up relationship; rather, it is flowing in both directions in a manner that both levels provide information to the other.

Conclusion

As demonstrated in this chapter, American foreign policy is not characterized by the direction of a single leader or a cohesive, uncontested national interest but rather a complex web of interactions among numerous actors pursuing the various missions and goals of their respective agencies. It is shaped and determined by several facets to include endogenous factors such as the presidency, Congress, the bureaucracy, and American public opinion as well as significant exogenous factors such as the Spanish-American War, World War II, the cold war, and the events on September 11, 2001. Although many have argued that the executive branch has *carte blanche* in the realm of foreign affairs, there are several constraints on its power through the delegation of powers in the U.S. Constitution as well as the numerous checks on presidential power by Congress. Furthermore, this chapter has displayed that there are other actors who have considerable influence and power in American foreign policy, such as Congress, the bureaucracy, and the American public. Overall, this chapter has demonstrated that American foreign policy has transformed over time through several influences and arguably will continue to evolve based on endogenous factors within the United States as well as exogenous influences in the international system.

References and Further Readings

- Allison, G. T. (1969). Conceptual models and the Cuban missile crisis. *American Political Science Review*, 63(3), 689-718.
- Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.
- Allison, G. T., & Zelikow, P. (1999). *Essence of decision: Explaining the Cuban missile crisis* (2nd ed.). New York: Addison Wesley.

- Almond, G. (1950). *The American people and foreign policy*. New York: Harcourt Brace.
- Ambrose, S. E., & Brinkley, D. G. (1997). *Rise to globalism: American foreign policy since 1938* (8th ed.). New York: Penguin.
- Carter, R. G., & Scott, J. M. (2009). *Choosing to lead: Understanding congressional foreign policy entrepreneurs*. Durham, NC: Duke University Press.
- Converse, P. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206–261). New York: Free Press.
- Edwards, G. C., III, & Wayne, S. J. (2006). *Presidential leadership: Politics and policy making* (7th ed.). Belmont, CA: Wadsworth.
- Entman, R. M. (2004). *Projections of power: Framing news, public opinion, and U.S. foreign policy*. Chicago: University of Chicago Press.
- Fisher, L. (2006). What is the appropriate role of Congress in national security policy? In G. C. Edwards III (Ed.), *Readings in presidential politics* (pp. 263–290). Belmont, CA: Wadsworth.
- Fletcher, S. R. (2000). *Global climate change: The Kyoto Protocol* (CRS Report for Congress No. RL30692). Washington, DC: Congressional Research Service.
- Garrison, J. A. (1999). *Games advisors play: Foreign policy in the Nixon and Carter administrations*. College Station: Texas A&M University Press.
- Garrison, J. A. (2001). Framing foreign policy alternatives in the inner circle: The president, his advisors, and the struggle for the arms control agenda. *Political Psychology*, 22(4), 775–807.
- George, A. L. (1980). *Presidential decisionmaking in foreign policy: The effective use of information and advice*. Boulder, CO: Westview Press.
- George, A. L., & George, J. L. (1998). *Presidential personality and performance*. Boulder, CO: Westview Press.
- Hart, P., Stern, E. K., & Sundelius, B. (1997). *Beyond group think: Political group dynamics and foreign policy making*. Ann Arbor: University of Michigan Press.
- Holsti, O. R. (1992). Public opinion and foreign policy: Challenges to the Almond Lippmann consensus. *Mershon series: Research programs and debates. International Studies Quarterly*, 36, 439–466.
- Holsti, O. R. (1996). *Public opinion and American foreign policy*. Ann Arbor: University of Michigan Press.
- Janis, I. (1983). *Groupthink: Psychological studies of policy decisions and fiascoes* (2nd ed.). Boston: Houghton Mifflin.
- Kennan, G. F. (1984). *American diplomacy* (Expanded ed.). Chicago: University of Chicago Press.
- Kingdon, J. (1995). *Agendas, alternatives, and public policies* (2nd ed.). New York: HarperCollins.
- Kissinger, H. (1994). *Diplomacy*. New York: Touchstone.
- Kowert, P. A. (2002). *Groupthink or deadlock: When do leaders learn from their advisors?* Albany: SUNY Press.
- Lippmann, W. (1925). *The phantom public*. New York: Harcourt Brace.
- Mingst, K. A. (2008). *Essentials of international relations* (4th ed.). New York: W. W. Norton.
- Papp, D. S., Johnson, L. K., & Endicott, J. (2005). *American foreign policy: History, politics, and policy*. New York: Longman.
- Peterson, P. E. (1994). The president's dominance in foreign policy making. *Political Research Quarterly*, 109, 215–234.
- Pika, J. A., & Maltese, J. A. (2008). *The politics of the presidency* (7th ed.). Washington, DC: CQ Press.

96

RACE, ETHNICITY, AND POLITICS

ANTHONY DeFOREST MOLINA

University of South Dakota

The field of racial and ethnic politics concerns itself with the role that race and ethnicity play in shaping the political behavior of individuals and groups, as well as the role that race and ethnicity play in how social, economic, and political institutions are constructed. This chapter provides an overview of that field by sketching out the major themes that exist in the body of scholarship known as racial and ethnic politics, as well as illustrating the manner in which the insights gained from such inquiry can be useful in understanding the political experience of minority groups in America. It begins with an overview of the basic terms, concepts, and political strategies that have been identified by scholars in the field. Next, a few examples are provided of the way in which these terms, concepts, and strategies may be applied by briefly examining the experience of some of the various minority groups in the American political system. Finally, a brief discussion of the future directions of the field is provided.

Racial and Ethnic Political Theory: Concepts and Strategies

In providing an overview of any field of inquiry, it is important to define the meaning of the basic terms and concepts that are used. As Michael LeMay (2000) has

pointed out, this is all the more true for the field of racial and ethnic politics because the terms used are often controversial and emotionally charged. In fact, the manner in which scholars in the field use terms such as *racism*, *prejudice*, and *discrimination* can sometimes be at odds with the manner in which these terms are used in the media and in popular discourse. Therefore, the following section of this chapter seeks to carefully define the core terms and concepts that are used in describing the experiences of minority groups in the American political system. It then turns to a brief overview of the various political strategies that ethnic and racial groups have utilized in coping with their status as a minority group.

Basic Terms and Concepts

Majority and Minority Status

The most basic of concepts in understanding the role of race and ethnicity in a political system are the concepts of majority and minority status. These two terms refer to the degree of power that groups have with respect to each other in a political system (LeMay, 2000). In other words, in any given political system, there will be a group that has the power to determine the norms and values of society, with the remaining groups being subject to those determined

norms and values. Thus, the term *majority* refers to the group in a political system that has the power to determine the norms and values by which social, political, and economic life is structured. Importantly, this group does not necessarily need to represent a numerical majority, as in the case, for example, of blacks in South Africa under apartheid rule. Although blacks in South Africa at this time represented an overwhelming numerical majority, they were subject to the official South African policy of apartheid (separateness), which included statutory requirements that excluded blacks from positions of power, denied them access to many public accommodations, and enforced a rigid social hierarchy in which white South Africans enjoyed a position of power and privilege. Under such conditions, white South Africans, though not a numerical majority, enjoyed superordinate status as a group in the political system and were therefore the majority group. In turn, the term *minority* in this context refers to those groups that are subject to the norms and values of the majority but, for one reason or another, do not enjoy membership in that group.

Ethnicity, Race, and Racism

Peter Rose (1964) has defined an ethnic group as a group whose members share a social and cultural heritage that is passed from generation to generation. According to Rose, this sense of belonging to the ethnic group can be deeply influenced by the manner in which the dominant group in society responds to the ethnic group. If the ethnic group is accepted by the majority group, they may come to identify themselves more closely with the majority group, and so in such cases the “sense of peoplehood” that they share as an ethnic group may weaken (LeMay, 2000). On the other hand, where the majority rejects the ethnic group, the bonds that hold the ethnic group together may actually strengthen as they seek to cope with their status as a minority group in the face of majority group hostility. Edgar Litt (1970) points out that these dynamics between groups are neither new nor unique to the American experience but are as old as the human condition itself. As he puts it,

The shared symbols, interests, affections, and real or imagined traits which draw some men together into the group or community are the walls which separate these men from others . . . for there to be “brothers” there must be “others.” (p. 4)

In contrast to *ethnicity*, the term *race* refers to the genetic makeup of individuals that accounts for the differences that we find in physical characteristics such as skin color, hair texture, and facial features. The term *racism* applies where individuals hold a belief that these differences in physical characteristics are directly linked to intellectual functioning, and on the basis of those differences, people make distinctions between races that are superior and races that are inferior (LeMay, 2000). Michael Banton (1967)

has argued that, from the standpoint of social science, it is the social dimension of race that is important in understanding the nature of race relations, rather than the biological dimensions. In other words, the focus of analysis for social science researchers should be related to how the concept of race is socially constructed and the implications that those constructions have for how individuals meet the expectations assigned to those roles.

Prejudice and Discrimination

The term *prejudice* refers to an attitude (usually a negative attitude) that is directed toward individuals who are perceived to belong to a particular group (Rose, 1964). Simply put, prejudice is the practice of prejudging an individual on the basis of his or her membership in a group, rather than on his or her unique individual characteristics. According to Kitano (1997), prejudice may develop for a variety of different reasons. For example, where one group dominates another for their own benefit, prejudice may develop as a means of keeping the other group in a subordinate position in order to justify the existing social order. Similarly, ignorance may serve as another source of prejudice because a lack of information about a racial or ethnic group can lead to the formation of stereotypes, which are oversimplistic and overexaggerated beliefs about a group. Selective perception and the projection of these images through the mass media and popular culture of society subsequently lead to their reinforcement. In contrast to prejudice, which, again, is an attitude, the term *discrimination* refers to behavior. Put otherwise, a person who acts on prejudicial views in dealing with members of a group is engaged in discrimination, or as LeMay (2000) calls it, “applied prejudice.” An employer, for example, may hold prejudicial views about a minority group without acting on those views. However, if the employer were to deny an individual employment because of prejudicial views about the group to which they are perceived to belong, then the employer would have engaged in discrimination.

Social Stratification and Segregation

The term *social stratification* refers to the process whereby individuals and groups are assigned to different roles and positions in society, which, in turn, results in an uneven distribution of the rewards and power that accompany those roles and positions (LeMay, 2000). A related concept, *segregation*, refers to the act of physically separating members of a minority group from the members of the majority group in order to limit contact between the two groups. Segregation may be either *de jure* (in law) or *de facto* (in fact). An example of *de jure* segregation can be found in the legal statutes or so-called Jim Crow laws of the post-Reconstruction Era South (Sitkoff, 1981). Such segregation codified in law the barring of African

Americans from access to various public accommodations such as hotels, restaurants, and public pools. In other cases, segregation statutes created an entirely separate set of social institutions on the basis of race, including public schools and military units. The U.S. Supreme Court upheld the constitutionality of these types of laws in its landmark ruling in *Plessey v. Ferguson* (1896) by establishing the separate but equal doctrine. Although the Court reversed itself almost 60 years later in its *Brown v. Board of Education* (1954) ruling by declaring that “separate is inherently unequal,” the practice of segregation proved difficult to undo. In lieu of de jure segregation, various manifestations of de facto segregation persist to this day (Barker, Jones, & Tate, 1999). In contrast to de jure segregation, de facto segregation occurs as a result of social practices that, although not codified in law, have a similar outcome in terms of maintaining the separation of a minority group from the majority group.

Acculturation and Assimilation

When different ethnic groups come into contact with one another, their members become conscious of the different norms and values that exist between them. As a result of contact between groups over time, acculturation takes place. Litt (1970) defines acculturation as “the process whereby the minority member absorbs the cultural ways, values, and lifestyles of the wider community, or that portion of the wider community which operates within the regional and class confines available to him” (p. 15). According to LeMay (2000), however, this process of acculturation is two directional, meaning that not only do minority groups begin to internalize norms and values from the majority group, but also—vice versa—majority group members begin to internalize some of the norms and values of the minority group. The decoration of Christmas trees, for example, now a common practice among majority group members in America, was absorbed via contact with German immigrants. Similarly, pizzas and quesadillas are now a mainstream staple of American cuisine as a result of contact with Italian and Mexican immigrants. Acculturation, then, simply refers to the process by which these sorts of practices make their way from one group to another.

Assimilation is a concept that is related to acculturation but refers to a much more deeply transformative process wherein an individual or individuals of one group begin to identify psychologically with the other group—that is, the minority group is essentially absorbed into the majority group as new members (LeMay, 2000). This is most often the case where majority group members are willing to accept the minority group (or individuals within that group) as members of their own group. In such cases, distinctions between the groups become less meaningful as the respective members begin to view themselves as all members of one group. In American society, for example, Germans were once considered a distinct minority group

in relation to the Anglo-American majority. However, as German immigrants began to adopt the English language and Anglo-American customs, the process of assimilation made the distinctions between German and Anglo-Americans less important, and over time they became assimilated into the majority group.

It is important to note here that not all minority groups can, or want to, assimilate into the majority group (Litt, 1970). For whatever differences existed between them and the majority Anglo group, German Americans were still Caucasian, primarily Christians, and shared a great many of the basic norms and values of the Anglo majority group. Thus, assimilation was relatively easy and desirable for them. Other groups, however, particularly those of African descent, have experienced a significant degree of acculturation but have at the same time been blocked from assimilation as a result of factors such as prejudice and racism (Rose, 1964). Thus, the distinction between acculturation and assimilation is an important one. As Litt (1970) explains, the minority group member “who undergoes acculturation is not necessarily freed from feelings of marginality or minority group status” (p. 15). In any event, minority groups often find that they must use some strategy, or combination of strategies, to cope with their status in the political system, and so this chapter now turns to an examination of the principal political strategies that minority groups use.

Minority Political Strategies

David Easton (1953) once famously defined politics as the authoritative allocation of values. Seen from this perspective, politics is a process whereby it is decided who gets what, when, and under what circumstances. In this process, minority groups are at a disadvantage in relation to the majority group in influencing the allocation of those values in a manner that benefits their group members. As noted previously, through the process of acculturation and assimilation, a minority group may seek acceptance into the majority group for the purpose of enjoying the privileges of majority group status. However, not all minority groups can or want to assimilate into the majority group. How minority groups are perceived by the majority, and whether they can or desire to assimilate into the majority, influences the choice of political strategies that minority groups employ to cope with their minority status. Following Litt’s (1970) taxonomy of minority political strategies, as well as the work of LeMay (2000), this section briefly discusses the three strategies that minority groups typically pursue: accommodationism, separatism, and radicalism.

Accommodationism

The term *accommodationism* refers to a political strategy in which the minority group (at least for the most part)

accepts the norms and values of the majority group and seeks to become part of that group (LeMay, 2000). Gaining entrance into majority group membership, in turn, allows its members to share in the privileges that accompany majority group status. LeMay has argued that in employing the strategy of accommodationism, minority groups will follow one of two routes: the economic route or the political route. In those instances where a minority group voluntarily migrates, possesses good job skills, and arrives at a time of economic prosperity, the economic route may be the most attractive form of accommodationism. In such cases, minority groups often rely on occupational niches and occupational queuing.

The term *occupational niche* refers to the presence of a high concentration of minority group members in an occupation in which the majority group does not feel threatened by their presence. In contrast to occupational niches, the term *occupational queuing* refers to the rank ordering of jobs in terms of their socioeconomic desirability. Minority groups that are willing to take less desirable jobs—low-skill jobs in agriculture or in meat processing facilities, for example—may be filled by minority group members without representing a threat to the majority. Successfully exploiting the use of occupational niches and occupational queuing allows minority groups to climb the socioeconomic ladder by providing improved educational and career opportunities for their children.

As a minority political strategy, however, economic accommodationism is not a viable option for all groups who may desire to pursue it. In some cases, minority groups find that the economic route is not open to them, as a result of poor economic conditions or as a result of discriminatory employment practices. The Irish, for example, on arriving in America were greeted with open hostility and employment discrimination—often unabashedly pronounced in signs that read “Irish Need Not Apply” (Rose, 1964, p. 33). In such instances, a minority group may choose to pursue a political route of accommodationism by exercising influence within the political system through the electoral process. To do so, however, the group must possess the means to organize themselves as a cohesive voting bloc that can influence the electoral process. Through a strategy of political accommodationism, then, a minority group may seek rewards from the ruling party in exchange for their votes. In those instances where a strategy of accommodationism is either unavailable or undesirable, the minority group may instead seek to pursue a strategy of separatism or a strategy of radicalism.

Separatism

Separatism is a political strategy that may be used by a minority group that rejects the norms and values of the majority group and consequently seeks to minimize contact with that group. According to Litt (1970), separatism occurs when a minority group “turns inward, creating its

own institutions to replenish social, psychological, and cultural values that cannot find fulfillment in the larger society” (p. 75). As with accommodationism, LeMay (2000) points out that groups who pursue separatist strategies may choose one of two routes. In the case of separatism, depending on the circumstances in which the group must operate, minority groups may opt for either a physical or psychological separation. The Mormons and the Amish are examples of minority groups that were able to achieve physical separation because of the availability of sparsely populated land in what was then the Western frontier of the United States. Other groups, however, such as the Hasidic Jews, or members of the Nation of Islam, unable to achieve physical separation because of their settlement in densely populated urban centers, were able to maintain psychological separation through the erection of semiautonomous social institutions in the midst of majority society.

It is worth noting here that separatism may be imposed on a minority group involuntarily. In such instances, separation does not represent a political strategy on the part of the minority, but rather, segregation (as discussed previously)—a social condition that is imposed—by the majority group. The term *ghetto*, for example, is rooted in the Italian word *borghetto*, which refers to the section of European cities in which Jews were once required to reside (Levin, 1968). Modern-day ghettos, which normally refer to blighted inner-city areas with a high proportion of minority group members, are constructed through the enforcement of discriminatory legal and social practices that restrict access to housing for minority group members to specific areas (Kusmer, 1978). These settlement patterns are distinctly different from those that result from separatism as a minority political strategy because they come about because of concerted efforts on the part of the majority group to disrupt the process of acculturation and assimilation of the minority group.

Radicalism

As opposed to accommodationism and separatism, radicalism represents a political strategy in which the minority group rejects the norms and values of the majority group and seeks to revise or replace those norms and values with their own (LeMay, 2000). As with accommodationism and separatism, radicalism can take shape in two particular ways. Litt (1970) refers to the first of these as old-style or ideological radicalism in which the ideological views of the minority group are significantly different than the majority, but the political behavior remains essentially consistent with accepted norms. Groups that used this approach, such as the Marxists of the late 19th and early 20th centuries, primarily relied on the electoral system in their efforts to win over the majority. In contrast, groups using the new-style approach to radicalism—as in the civil rights groups of the mid-20th century, for

example—shared an ideological position that was not necessarily contrary to prevailing majority group values. What distinguishes new-style radicalism from old-style radicalism, however, is the use of political behavior that is very much at odds with accepted norms of political behavior. According to Litt, these radical political behaviors encompass a wide range of tactics from nonviolent direct-action protest to full-scale revolution and flow from a sense of urgency and passion. He writes, “The new radicalism is emotional rather than ideological . . . it is based on a preference for action rather than for cognition and polemics” (p. 101).

Applications: The Political Experience of Minority Groups in America

In the space available here, it is not possible to cover the diverse variety of experiences that ethnic and racial minority groups have had in the American political system. The United States is a land of immigrants from every part of the world, and as noted above, different groups have met with varying degrees of acceptance and resistance as they attempted to carve out their niche in American society. Nevertheless, to the extent that such an exercise can highlight the manner in which the foregoing theoretical concepts can be applied, it is helpful to examine certain aspects of the minority group experience. Still, the following overview of selected minority group experiences in the United States is by no means intended to be an exhaustive account of those experiences. Rather, the experiences roughly sketched out below are simply offered as examples of the manner in which some groups have sought to cope with their minority status and to help illustrate the extent to which the theoretical concepts associated with racial and ethnic politics can be usefully applied in an analysis of minority group political behavior.

Irish Americans

Beginning in the 1840s, Irish immigration to America began to increase significantly due in large part to political repression that they experienced under British rule, but also by periodic famines in which many thousands of Irish starved to death (Harrigan & Vogel, 2000). These immigrants, who arrived on American shores with relatively small financial resources and few job skills, were viewed with suspicion predominantly because of their Roman Catholic faith (Rose, 1964). The dramatically increasing number of Irish immigrants alarmed many members of majority American society, and so during this period an anti-Irish, anti-Catholic backlash took place. According to Harrigan and Vogel (2000), a nativist movement arose to protect American society from this threat; churches and convents were burned by angry mobs; publications appeared such as Maria Monk's anti-Catholic book, *Awful Disclosures*; and a political party

called the Know-Nothings formed to protect the “integrity of American society” (p. 62). In their efforts to weed out popery, the Know-Nothings' stated platform included restrictions on Irish immigration and the barring of persons of Irish descent from holding office.

Harrigan and Vogel (2000) argue that the Irish reacted to this hostility by taking the characteristics that were most despised by majority American society and embracing those characteristics as virtues rather than vices. For the Irish, chief among these characteristics was the Roman Catholic religion. Catholicism gave the Irish a sense of uniqueness and bound them together as a group. This, in turn, became an effective tool of political mobilization since it enabled the Irish to build electoral coalitions across class lines. Furthermore, the Catholic Church recognized that if the Catholic faith were to be preserved in this environment, it would be necessary to embark on substantial development of parish schools and churches. Harrigan and Vogel explain that this decision had two important consequences. First, it organized the Irish into relatively discreet geographical units, and second, it created a great deal of construction work that could be channeled to Irish construction companies and laborers. Thus, the parish became the center of Irish social life in America, and the activities of the parish and the work of the construction projects led to the emergence of an institutional framework that could serve the needs of the Irish as a minority group. Furthermore, because the Irish were geographically bound together as parish groups by the church, they were subsequently able to emerge as a very influential voting bloc that eventually gained control of the Democratic Party in major cities where they resided (LeMay, 2000). Over time, this led to Irish control of the political machines in those cities.

The Irish, of course, did not invent machine politics. What they were able to do, however, was to take over pre-existing Democratic Party organizations established during the Jacksonian era (Harrigan & Vogel, 2000). Control of these political machines allowed the Irish to secure patronage jobs and to award contracts. Patronage opened up important avenues for the Irish in their advancement up the American socioeconomic ladder since jobs in the bureaucracies of the cities that they controlled went to their group members. Thus, in places like New York City and Boston, the Irish were disproportionately employed in public service positions such as the fire and police departments. Likewise, the ability to award contracts multiplied the success given to them by the church in construction and real estate. As a result, the Irish were able to successfully use accommodationism to secure their position in the American political landscape. Successive groups of immigrants have used this strategy as well, but as LeMay (2000) points out, “they used it less successfully because of changing conditions and because the Irish, who arrived before them, were more reluctant to budge from their newly acquired middle-class ‘rung’ than were the Yankees before them” (p. 115).

Mexican Americans

Mexican Americans, or Chicanos, consists of two distinct subgroups: Spanish Americans and Mexican immigrants. Rose (1964) explains that Spanish Americans became American citizens by default when the lands that they occupied for over 4 centuries in the Southwest were ceded to the United States at the conclusion of the Mexican American War in 1848. Gilberto Lopez y Rivas (1973) has argued, however, that there is a common misconception among Americans that the people in these “conquered territories received Anglo rule with open arms” (p. 9). Rather, he points out that a long history of resistance and struggle characterizes their experience in the American political system. In contrast to the Spanish Americans of the Southwest, Mexican immigrants began to come to the United States from Mexico in large numbers beginning in the 20th century. Although some left because of the turmoil that resulted from the Mexican revolution, many more were attracted by economic opportunities.

According to LeMay (2000), the rate and degree of acculturation and assimilation among Mexican Americans is very low. This is due in large measure to their close proximity to Mexico, which means that their culture is able to survive more easily than is the case with other groups. Additionally, Mexican Americans continue to have a strong sense of pride in the language and cultural heritage of Mexico. As a result of these and other factors, Mexican Americans have had relatively limited involvement in politics for much of their history in the United States. Some early groups did begin to emerge in the 20th century, such as the Order of the Sons of America and the League of United Latin American Citizens, both of which were accommodationist in orientation (LeMay, 2000). Following World War II, however, organizations that were much more politically oriented began to emerge in an effort to increase voter turnout, endorse candidates, and engage in community organizing (Garcia & Garza, 1977).

Beginning in the 1970s, it became much more common for Mexican Americans to employ radicalism as a political strategy. These groups, deeply influenced by a sense of cultural pride, were more militant than their predecessors. They included groups such as the Brown Berets, the Crusade for Justice, and various student organizations. The most widely recognized group of this movement was *La Raza Unida* (the United Race). According to F. Chris Garcia and Rudolph O. de la Garza (1977), as opposed to accommodationists who sought to “work within the system,” radical Chicano groups such as these were “willing and ready to resort to tactics which violate[d] existing political norms in order to attain their goal of restructuring American society” (p. 43). They caution, however, that distinguishing the two groups in this manner “grievously distorts the relationship” shared by them and that the “differences between them are less important than the similarities that [bound] them” (p. 43).

As the Mexican American population has grown and participation rates have gone up, its political influence in American politics has increased accordingly (Browning, Marshall, & Tabb, 1984). Rodney Hero (1997), for example, argues that in the city of Denver, where Hispanics accounted for only 20% of the population, they were nonetheless able to elect Federico Peña by building an electoral coalition that crossed class and racial lines. In turn, Peña was successful at bringing Mexican Americans into city government to a level that is roughly proportionate to their size in the population.

African Americans

As a minority group in American society, the experience of African Americans is unique because they represent the only group in American society to have arrived as a result of slavery. With some exceptions, most of the African Americans residing in the United States today are the descendents of Africans who survived the experience of slavery, as well as the subsequent decades of hostility and discrimination that they faced at the hands of the majority group. As Cornell West (1999) has pointed out, “Slavery is nearly as old as human civilization itself,” but “the distinctive feature of New World [American] slavery was its ‘racial’ character” (p. 51). As a result, the discrimination directed toward African Americans went well beyond their status as slaves, but was targeted directly at their racial identity. Thus, following their emancipation from slavery, with few exceptions, the struggle of African Americans for social, economic, and political equality continued largely unabated.

Early on in this struggle, two styles emerged around the persons of Booker T. Washington and W. E. B. Du Bois. Washington argued for an approach based largely on the principles of economic accommodation. Focusing on individual achievement, Washington called for African Americans to integrate into mainstream society by acquiring basic job skills and downplayed the role of political activity in pursuit of civil rights. In contrast to Washington’s approach, Du Bois argued that African Americans would need to emerge as a group if they hoped to break down the barriers that they faced. In 1905, Du Bois helped found the Niagara Movement, which quickly died out but whose cause was soon taken up by the National Association for the Advancement of Colored People (NAACP). Over time, the NAACP has sought to fight racial discrimination primarily through courtroom litigation, with its most significant victory being the landmark *Brown v. Board of Education* (1954) school desegregation case.

African Americans and the Strategy of Separatism

Marcus Garvey’s Black Nationalism Movement of the 1920s is an example of an African American separatist movement (LeMay, 2000). Founder of the Universal

Negro Improvement and Conservation Association, Garvey argued that people of African descent needed to establish their own set of social institutions, including banks, schools, and companies, and to ultimately reclaim the African continent for people of African descent. Drawing on Booker T. Washington's philosophy of economic independence (but rejecting his individualist approach), Garvey sought to connect all of the black peoples of the world for commercial and industrial exchange through the establishment of a shipping company, the Black Star Steamship Line. However, because of a series of financial scandals involving the Black Star Steamship Line, he began to lose credibility among mainstream blacks. A conviction for tax evasion and mail fraud resulted in a prison sentence from 1925 to 1927, and on his release he was deported from the United States. Garvey died in 1940 at the age of 52 in West Kensington, England. Since his death, however, his role as an early advocate for black nationalism became more deeply appreciated, particularly by leaders of the black nationalist movements of the 1960s such as Malcolm X and the Nation of Islam.

African Americans and the Strategy of Radicalism

In December of 1955, Rosa Parks of Montgomery, Alabama, was arrested and jailed for refusing to obey a bus driver's order that she give up her seat to a white man. With this act of defiance, Parks is credited with sparking the now famous Montgomery bus boycott and the modern civil rights movement. It was through the attention that the boycott received that Martin Luther King Jr. was first thrust into the national spotlight. Having recently arrived in Montgomery as the young new pastor of the Dexter Avenue Baptist Church, he and other pastors formed the Montgomery Improvement Association and later the Southern Christian Leadership Council, of which he was elected president in 1957.

Drawing on the philosophy of Mahatma Gandhi, King showed through his efforts that nonviolent direct action could be combined with legal tactics to successfully fight racial discrimination in the United States. In later years, organizations such as the Southern Christian Leadership Council and the Student Non-Violent Coordinating Committee engaged in a combination of legal battles along with direct action tactics that eventually led to the Civil Rights Act of 1964. In April 1963, King was in Birmingham, Alabama, where he had been asked to come in support of the campaign of nonviolent resistance that was taking place there. During the course of those events, he was arrested and placed in solitary confinement. On April 12, 1963, a newspaper ad was taken out by eight clergymen calling for an end to the demonstrations and criticizing King for his role. In his "Letter From Birmingham Jail," King responded to these criticisms and summarized the ideas behind the nonviolent direction action strategy that he advocated. According to King (1998),

Nonviolent direct action seeks to create such a crisis and foster such a tension that a community which has constantly refused to negotiate is forced to confront the issue. It seeks to dramatize the issue so that it can no longer be ignored. (p. 190)

Shortly after the events of Birmingham, on August 28, King addressed a crowd of 250,000 people who had gathered in front of the Lincoln Memorial in Washington, D.C., and delivered his famous "I Have a Dream" speech. The following year saw passage of the landmark 1964 Civil Rights Act, which, among other things, prohibited discrimination in public accommodations, required desegregation of public schools and colleges, and prohibited private sector employers from discriminating in employment on the basis of race, religion, sex, or national origin. Subsequently, the Voting Rights Act of 1965 provided a strong legal framework to reverse historical patterns of African American voter disenfranchisement.

The 1964 Civil Rights Act and the 1965 Voting Rights Act were significant milestones in the African American struggle for political equality. However, it has been observed that although by the end of the 1960s the legal barriers that had prevented African Americans from fully participating in American social, political, and economic life had been removed, African Americans remained at "the lower ends of just about every segment of American life" (Barker, Jones, & Tate, 1999, p. 26). In fact, the struggle for full political incorporation is likely to continue. To be sure, there are many examples of African Americans who have met with great success. In addition to the many successful African American business men and women, artists, scientists, and elected officials, Colin Powell was appointed the first African American Secretary of State in 2001, and he was followed in that office by Condoleezza Rice, an African American woman. Of even greater consequence, of course, Barack Obama, Senator from Illinois, the son of a Kenyan goat herder, was elected president of the United States in 2008. However, as Barker et al. (1999) have argued, "Individual successes, no matter how spectacular, do not change the reality of group oppression" (p. 350), nor does it erase its persistent legacy. All the same, great progress has been made, and despite the fact that they are likely to remain salient features of American political life, the dynamics of race, ethnicity, and politics is evolving in new directions. It is to a consideration of those future directions that this chapter now turns.

Future Directions

With the election of Barack Obama as the 44th president of the United States, it is common to hear speculation in the popular media that America has entered a new post-racial political era. Indeed, beginning as early as the late 1980s, some African American candidates for political

office found electoral success by playing down racial issues and seeking broad support from white voters. This campaign strategy of deracialization has been described by Huey Perry (1996) as a useful analytical tool for understanding African American politics since the presidential administration of Ronald Reagan. According to Perry, Reagan was able to considerably strengthen the conservative mood of national politics at that time, and this mood included, in part, an implicit desire for African Americans to soften their desire for full incorporation into American political and economic life. Thus, in the decades that followed, African American candidates seeking election in majority white electoral districts found some value in deemphasizing issues related to race. Of course, Obama's election to the presidency could not have been possible without the widespread support of white voters, and so, at least to some extent, deracialization remains an important concept in understanding the future of American politics.

Nevertheless, it is not at all clear that the election of America's first African American president indicates that the United States has entered a postracial era in which questions of race become entirely moot. Howard Winant (2004) argues that the very social, political, and economic structures within which Americans operate have been polluted by racism on a global scale. It is his assessment that because racial domination has been so successful in shaping social, political, and economic life over the past few centuries, only radical upheavals can result in real change. In that vein, he points to the U.S. Civil War and World War II as seminal moments of upheaval where progress toward racial equality (however small) was able to take place. Although racial and ethnic discrimination may no longer be sanctioned by law, its legacy persists insofar as its effects are still plainly visible in the significant gaps that exist between the socioeconomic well-being of whites on the one hand and racial and other ethnic minorities on the other. Thus, race and ethnicity shall continue to play an important role in American political life and also remain an important area of scholarly inquiry among political scientists.

In that vein, Julie Novkov (2008) has described the emerging body of scholarship within the subfield of racial and ethnic politics. According to Novkov, much of the recent research in racial politics has incorporated the tools of American political development to illuminate the manner in which race has played a critical role in the development of American political institutions. Such scholarship seeks to highlight the degree to which race as a cultural concept influenced how American social institutions were designed and developed. Additionally, research in this body of work has emphasized "how race interacts with the fundamental ideological commitments and shifts in American politics," as well as how "current institutions continue to reflect and refract this history of racialized struggle" (Novkov, p. 651).

Conclusion

This chapter has provided an overview of scholarship in the discipline of political science that focuses on the role that race and ethnicity play in the American political system. Such scholarship is concerned with not only the role that these social constructs play in shaping the behavior of individuals and groups, but also the role that they play in how social, economic, and political institutions are constructed. The chapter's purpose was to discuss the major themes that exist in that body of scholarship and to highlight some of the ways in which the insights gained from it can help us to understand the political experience of minority groups in America. The chapter began with an overview of the basic terms, concepts, and political strategies that are associated with racial and ethnic politics as a scholarly endeavor and then turned to a few examples of the way in which these terms, concepts, and strategies may be applied. Finally, a brief discussion of the future directions of the field suggested that, although significant gains have been made by racial and ethnic groups, race and ethnicity are likely to remain salient features of American political life for the foreseeable future.

References and Further Readings

- Banton, M. (1967). *Race relations*. New York: Basic Books.
- Banton, M. (1987). *Racial theories*. Cambridge, UK: Cambridge University Press.
- Barker, L., Jones, M., & Tate, K. (1999). *African Americans and the American political system* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Browning, R., Marshall, D. R., & Tabb, D. (1984). *Protest is not enough: The struggle of blacks and Hispanics for equality in urban politics*. Berkeley: University of California Press.
- Easton, D. (1953). *The political system: An inquiry into the state of political science*. New York: Knopf.
- Garcia, C. F., & Garza, R. O. (1977). *The Chicano political experience: Three perspectives*. North Scituate, MA: Duxbury Press.
- Harrigan, J., & Vogel, R. (2000). *Political change in the metropolises*. New York: Longman.
- Harris Lacewell, M. (2004). *Barbershops, bibles, and BET: Everyday talk and black political thought*. Princeton, NJ: Princeton University Press.
- Hero, R. (1997). Latinos and politics in Denver and Pueblo, Colorado: Differences, explanations, and the steady state of the struggle for equality. In R. Browning, D. R. Marshall, & D. Tabb, *Racial politics in American cities* (pp. 247–258). New York: Longman.
- Judd, D., & Kantor, P. (2002). *The politics of urban America: A reader*. New York: Longman.
- King, M. L., Jr. (1967). *Where do we go from here: Chaos or community*. New York: Bantam.
- King, M. L., Jr. (1998). *The autobiography of Martin Luther King, Jr.* (C. Carson, Ed.). New York: Warner.

- Kitano, H. (1997). *Race relations* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Kusmer, K. (1978). *A ghetto takes shape: Black Cleveland, 1870-1930*. Chicago: University of Illinois Press.
- LeMay, M. (2000). *The perennial struggle: Race, ethnicity, and minority group politics in the United States*. Upper Saddle River, NJ: Prentice Hall.
- Levin, N. (1968). *The Holocaust: The destruction of European Jewry 1933-1945*. New York: Thomas Crowell.
- Litt, E. (1970). *Beyond pluralism: Ethnic politics in America*. Glenview, IL: Scott, Foresman.
- Lopez y Rivas, L. (1973). *The Chicanos: Life and struggles of the Mexican minority in the United States*. New York: Monthly Review Press.
- Novkov, J. (2008). Rethinking race in American politics. *Political Research Quarterly*, 61(4), 649-659.
- Perry, H. (1996). An analysis of major themes in the concept of deracialization. In H. Perry (Ed.), *Race, politics, and governance in the United States* (pp. 1-11). Gainesville: University Press of Florida.
- Plessey v. Ferguson, 163 U.S. 537 (1896).
- Rose, P. (1964). *They and we: Racial and ethnic relations in the United States*. New York: Random House.
- Sitkoff, H. (1981). *The struggle for black equality, 1954-1980*. New York: Hill & Wang.
- U.S. Riot Commission. (1968). *Report of the National Advisory Commission on civil disorders*. New York: Bantam.
- West, C. (1999). *The Cornell West reader*. New York: Basic Civitas Books.
- Winant, H. (2004). *The new politics of race: Globalism, difference, justice*. Minneapolis: University of Minnesota Press.

97

GENDER AND POLITICS IN THE UNITED STATES

ANGELA L. BOS

College of Wooster

PAULA L. O'LOUGHLIN

University of Minnesota, Morris

In the last few decades, women in the United States have made great strides in politics. Although women have historically voted in lower numbers than men, a higher percentage of women have registered and voted in presidential elections than men since 1984. Women now also win election at rates comparable to their male counterparts. In Congress, women have made substantive policy changes that positively influence women. Beyond Congress, women have achieved other political successes. Hillary Clinton almost gained the Democratic presidential nomination in 2008, Sarah Palin was the second woman to be a major party vice-presidential nominee, and Condoleezza Rice recently served as the first black woman secretary of state.

The successes of women in politics raise a few important questions. First, is the political glass ceiling broken? Do women still face barriers in participating in politics based on their gender? If there are barriers, what are they and how can they best be minimized? And second, when women engage in politics—whether as participants in local city council meetings, as voters, or as members of congress—does their behavior make a difference? Do women have distinct political preferences from men? And if so, what explains this? Finally, what potential value lies in more women engaging in politics? This chapter proceeds by first introducing the most important questions and then reviewing relevant work addressing each question in the areas of women as

political actors, women as political candidates, and women in political institutions.

Theory and Literature Review

Women as Political Actors

When considering the role of gender in U.S. politics, it is first important to consider how gender influences citizens' perception of politics and their role in it.

Political Attitudes

Whether men and women have distinct political preferences and, if so, why these differences exist has been a primary research focus. Women report some policy preferences distinct from men. Women are less supportive of the use of military force in areas of foreign policy and are more liberal in their desire for the government to provide services, jobs, and health care to citizens. The most prevalent and studied attitude difference between men and women is the gender gap: differences between men and women in both party ideology and vote choice, with women being more likely than men to identify with the Democratic party and to support liberal, Democratic candidates. The gender gap has existed at least since 1964 (Kaufmann & Petrocik, 1999), but

disappeared briefly post–September 11, 2001, only to reemerge in 2004.

Scholars' investigation of the causes for these differences offers several possibilities on the origins of the gender gap. The first attitudes theory adopts the view that underlying issue preferences, particularly on social issues, explain the gap. This is supported by work that suggests feminism has contributed to the gender gap by promoting a so-called women's perspective that influences policy and promotes an ideology of equality and sympathy for the disadvantaged. This theory also comports with the idea that since women are more likely than men to identify with certain issues, the party that performs better on these issues will be more likely to have larger numbers of female voters (e.g., Kaufmann & Petrocik, 1999).

Although greater numbers of women have voted Democratic than Republican, in terms of issue advocacy, Sanbonmatsu (2002) finds that the Republican and Democratic Parties are largely similar on most gender issues. She argues that despite few compelling policy differences, Democrats are more willing to support and discuss gender issues. Abortion is the only issue that has become strongly partisan, which partly explains why feminists ally much more closely with the Democratic Party than the Republican Party. Overall, however, neither party has made gender issues a significant part of their party platform.

A second salience hypothesis posits that perhaps men and women weigh issues differently when evaluating parties and candidates. The idea is that the gender gap does not stem solely from different political orientations, but rather men and women might find different issues salient in their political decisions. Kauffman and Petrocik (1999) test the salience model alongside the attitude model and find evidence of the attitudes hypothesis in 1992 and the salience hypothesis in 1996. Both attitudes and salience explain the gender gap to some degree, but the context of the election influences how voters weigh various criteria.

A third explanation for the gender gap suggests it originates from changing male attitudes. Simply, there is an ongoing defection of party members from the Democratic Party to the Republican Party. The gap exists, at least partially, because men are becoming Republican far faster than women.

Research indicates contextual variables shape the gender gap's varying size and causes. Although women tend to be more likely to vote Democratic than men are, the candidates or issues influence the size of the gap. This theme is echoed by other scholars like Kauffman and Petrocik (1999), who state that the gender gap is responsive to political, economic, and social factors. Specifically, the gap increases as the political climate becomes more conservative, when the economy is poor, and as women as a group become more economically vulnerable and government programs are threatened. Only with issues of use of force and compassion does the size of the gender gap stay constant (Norrande, 1997).

Political Participation

Research on gender and political participation has examined the extent to which women participate in a broad range of political acts—from conventional to unconventional—as compared to men and to explain these differences. In many areas—such as voting, participating in local politics, and protesting—women participate largely in equal numbers as men. In a few areas, particularly in campaign contributions, belonging to political organizations, and in contacting public officials, women participate less.

Scholars have tried to explain these small but persistent gender differences in political participation in multiple ways (Burns, Schlozman, & Verba, 2001). Women are generally disadvantaged when it comes to the resources—financial and otherwise—that facilitate political activity. Scholars conclude that if women had more resources, their overall political activity levels would be very close to men's. Gender differences in educational attainment, although negligible among younger generations, continue to give men relatively more civic skills and networks to promote civic engagement. In combination with differences in educational attainment, workplace segregation means men are more likely to have jobs that cultivate the skills necessary to support political participation. And although women are more involved than men in certain types of organizations—for example, religious organizations—men in these organizations are more likely than women members to take on leadership positions that help them acquire civic skills (Burns et al., 2001).

Beyond political participation, several other gender gaps in political attitudes exist. Women tend to be less politically interested, informed, and efficacious than men, and these factors partially explain the gap between women and men in political activity. Women know less about politics based on standard political knowledge batteries as well, but half of the difference is explained by women being less likely to guess when presented with a question they are unsure how to answer (Mondak & Anderson, 2004).

Women as Political Candidates

Descriptive Representation

When one thinks about women as candidates, one might wonder whether it matters if women are elected to political office. Theories of democracy present multiple ways in which the interests of groups are represented in government. Descriptive representation is achieved by our governmental bodies reflecting the composition of our society. With regard to gender, pure descriptive representation would require that the percentage of women in political bodies reflect their relative percentage in the population. In contrast, substantive representation requires that the interests of women (or other groups) be raised in

government. Although many argue that male representatives could adequately raise the substantive interests of women, research finds numerous benefits to descriptive representation, the most compelling that when group members represent themselves, the overall substantive representation of their group interests improves (Mansbridge, 1999). Several other advantages have emerged as well. For example, the presence of women candidates increases women's civic engagement (Burns et al., 2001), and in districts where women are elected, female constituents are more likely to participate in politics and to have greater senses of political efficacy and political competence (High Pippert & Comer, 1998).

Explaining Women's Descriptive Underrepresentation

Since women compose approximately 50.9% of the population but hold just under 17% of seats in the current 111th Congress, they are descriptively underrepresented. The question that scholars have asked is why. Many argue that it could be largely explained by the so-called institutional inertia created by our incumbency-driven system, which creates very few openings for women to seek. Recent research, however, finds that term limits that are meant to address this have not had as large a positive effect as anticipated (see Bernstein & Chadha, 2003). Cross-national research finds that proportional representation systems, over winner-take-all systems like that in the United States, lead to more women in office. In the United States one sees a similar effect since one sees more women in office in states with multimember districts for state legislative seats. A third explanation begins with the idea that there are not enough qualified women in the pipeline to pursue political office. The argument proceeds that if more women pursued careers in the fields that most often precede political office—those in education, law, business, and politics—more women would seek and win positions. Although this is likely to have some effect on the number of women in office, simply having more women in the pipeline will not necessarily increase the number of women pursuing office since women are far less likely to perceive themselves as qualified to run for office, even when they hold similar qualifications to men (Lawless & Fox, 2005).

Because of these gender differences in self-perception among qualified candidates, outside encouragement becomes essential to increasing the number of women pursuing elected office. Women are twice as likely as men to run for office when encouraged by others, such as party leaders, to do so (Lawless & Fox, 2005). However, women are far less likely to receive encouragement from party leaders than men, and a majority of women candidates running for local office report they were actively discouraged by political party leaders (Niven, 1998). Prospective women candidates also perceive it is more difficult to be elected and to raise money for a campaign than their male counterparts (Lawless, 2009). Thus, self-perceptions and

differential political ambition among qualified men and women candidates seems to partially explain women's underrepresentation, especially when one considers that women are less likely to be encouraged to run but more reliant on such encouragement.

Voters Evaluating Women Candidates

The stereotypes voters have of women and how these are at odds with their expectations of so-called good politicians are also important for understanding women's descriptive underrepresentation. There is no evidence that gender stereotypes directly impact vote choice or electoral outcomes, but it is clear that stereotypes guide voters' evaluations of women candidates with regard to their beliefs, traits, and issue strengths. Compared with their male counterparts, women candidates are seen as more liberal (Koch, 2000), as possessing feminine (empathy, compassion) over masculine (assertive, competence) traits, and as capable to handle feminine-related issues (education, poverty) over masculine issues (military crises, foreign policy) (Huddy & Terkildsen, 1993). Thus, the stereotypes of women politicians do not match the male traits voters find important for those in all political offices. However, some positive stereotypes—for example, that women candidates are seen as agents of change—may benefit women candidates in certain electoral contexts.

The role of stereotypes is context dependent. For example, voters are more likely to rely on gender cues as information shortcuts in low-information contexts (McDermott, 1997). Stereotypes can be more or less helpful to women candidates based on the level of office they seek. The perceived traits and issue competencies of women seeking statewide offices such as governor—as opposed to U.S. senator—are more similar to the traits and issues voters desire for someone elected (Kahn, 1996). The broader issue context also matters. For example, the issue context in 1992, the so-called year of the woman, focused on gender issues, and thus, gender-related considerations were more important in the decisions voters made (Dolan, 2001). Lawless (2004) finds that in the post–September 11, 2001, atmosphere, stereotypes work against women candidates because citizens deem men more capable to legislate on issues such as military crises and national security.

Some scholars find evidence that women are more likely than men to want to see women in office and are more likely to vote for women candidates (Plutzer & Zipp, 1996). Research indicates that this support is not based on in-group preference but because they are concerned underrepresentation will leave certain issues overlooked (Paolino, 1995). Further, women with stronger identities as women, feminists, or as liberals are more likely to prefer female representation (Rosenthal, 1995). Recent research, however, refutes the conventional wisdom that women voters are more likely to support women candidates than men (Dolan, 2008).

Gender and Electoral Campaigns

Women win political races at rates comparable to their male counterparts and in similar types of races campaign fund-raising receipts are comparable. However, women candidates of all types are still more likely to face gender-based challenges in media coverage, voter stereotypes, and campaign strategy.

Where women run for office matters. Based on a variety of partisan, ideological, geographic, racial, and socioeconomic factors, a small number of congressional districts (18) are considered woman-candidate-friendly, and many more (153) are not woman-candidate-friendly (Palmer & Simon, 2006). Second, women incumbents are more likely to be challenged in primaries and to have candidates compete to run against them in the opposing party's primary (Palmer & Simon).

The mass media does not always treat female candidates equitably. Compared to their male counterparts, women are more likely to receive less coverage overall, more unfavorable news coverage including more horse race coverage, and less coverage focused on their issue agendas (Kahn, 1996). The level of disparity depends on the level of office: Press coverage is more favorable to female gubernatorial candidates than to U.S. Senate candidates.

Women and Political Parties

Scholars have also considered whether political parties enable women to gain political office or rather serve as gatekeepers to them. In U.S. political parties, women have almost managed to achieve gender parity but still hold few leadership roles. A large comparative politics literature finds that stronger, more centralized, and institutionalized political parties increase women's representation (see Caul, 1999). There is a dearth of research on the role of political parties in women's underrepresentation because U.S. political parties are comparatively weak. Burrell (1993) suggests political parties are no longer negative gatekeepers for women candidates, but neither do they control the nomination process that could facilitate the nomination of more women candidates. Sanbonmatsu (2006) finds that fewer women run for and hold elected office where parties are more likely to engage in gate-keeping activities.

Women in Political Institutions

Women in Legislatures

The first stream of research on women in legislatures examines whether gender affects the policy priorities of legislators. The initial work on this question suggests women's presence in the legislative arena has a discernible effect. In contrast to male legislators, women legislators see bills that focus on women, children, and families as most important and are more likely to invest their legislative capital in these areas.

If women representatives have distinct policy agendas, does their legislative behavior also differ from their male counterparts? The most recent research suggests gender's influence occurs from bill sponsorship and agenda setting to floor votes. Male and female members of Congress show statistically significant differences in the types of bills they sponsor and cosponsor, support in committee, and support on the floor (Swers, 2002). Women are more likely to sponsor or cosponsor legislation dealing with education, child rights, civil rights, economic equality legislation, women's health and welfare, and women's issues. However, although women members of Congress will also vote across party lines on women's issues, there are many behavioral differences among women. Gender's effects are mediated by political contextual factors such as partisanship, the partisan balance of power, the degree of polarization, and representatives' previous voting records. Increasing partisan polarization has made it more difficult to cross party lines on gender issues. In addition, Swers found that Democratic women's likelihood of speaking up on gender issues increased when they were the minority party while Republican women representatives' did not.

A third wave of research considers whether women legislators' style of work varies from that of their male counterparts. Generally, this has been found to be true. (Kathlene, 1994; Rosenthal, 2000). Kathlene (1994) finds that women legislators act differently than their male colleagues in committee hearings. Specifically, women legislators are less likely to speak on legislation regardless of their seniority, potential bill sponsorship, or party. Even when an issue is a high priority for a woman legislator or she is a committee chair, she is less likely than her male colleagues to speak in committee hearings. A study of women as committee chairs suggests other differences in behavior in this role. For example, women committee chairs are more consensus oriented than their male colleagues. However, institutional norms are more influential than gender.

Overall, the extent to which gender influences legislative outcomes, however, may be dependent on the number of women in the legislature and whether a critical mass of women members is achieved. In terms of legislative priorities, women legislators' distinctive goals become more pronounced when they have the strength of numbers. Women legislators' likelihood of success in passing legislation increases as the number of women in the legislature grows (Thomas, 1994). Vega and Firestone (1995) argue that when there are not very many women members, the effectiveness that can be tied to the issues that women—as a cohesive group—think are important decreases. As a result, those issues lose institutional support. In contrast, when women representatives are cohesive and work together to structure the congressional agenda, their capacity to influence it becomes stronger. Saint-Germain (1989) also concludes that the amount of time a female member will spend on issues that are important to her as a woman is at least partly a function of the proportion of women in

the chamber. She argues that the fewer the number of women in the House, the less likely they are to be effective in addressing women's issues. Other work suggests that a critical mass of women legislators is less important for women to sponsor legislation on women's issues (Bratton, 2005).

Recent research examining internal institutional norms and gender suggests why, in terms of legislative outcomes, gender remains fairly insignificant. Women representatives' historically low levels of formal institutional power have resulted in gender's minimal effects on legislative policy. Although women are capable of earning seniority and are as legislatively effective as their male colleagues, they have not been as effective because there have been fewer of them. Compared to their male colleagues, women members of Congress have historically lacked the seniority to chair prestige committees or hold elected office in the party hierarchy. Thus, they have been less able to influence legislation. Any gender differences in legislative preferences that might exist thus have been neutralized. Swers's (2002) study of gender and congressional agenda setting also suggests women often lack the institutional influence to push their preferred legislation.

In brief, previous research clearly indicates legislators' gender is significant. Gender interacts with representative roles, institutional influence, and bill sponsorship directly. However, its influence on women representatives is clearly contingent on other forces.

Women in the Courts

Research on gender and the courts has focused on two primary questions: the factors that affect women's likelihood of judicial selection and whether gender shapes judicial decision making. Studies of judicial selection are mixed as to what is most likely to lead to gender diversity on the bench. Some work suggests that how judges are picked matters. This research indicates women are more likely to be on courts when selection is done by appointment, not elections (Williams, 2007). However, others find that the larger the number of seats on a court, the more chance a woman will be selected for it.

The verdict on whether and how gender affects judicial decisions is mixed. Early work that looked just at male and female judges showed no significant gender differences in sentencing though male judges were inclined to be more lenient to women defendants (Gruhl, Spohn, & Welch, 1981). Scholars have also considered whether women judges use different models of legal reasoning. Research on this question suggests, however, that judges use the same legal reasoning irrespective of gender (Allen & Wall, 1993).

Other work on the relationship of gender to judicial sentencing suggests gender's effects on judicial decision making are more complex. When a judge's gender is considered along with other factors, such as judicial region as

a predictor of judicial choices, it is no more important than other characteristics. At the same time, however, the type of cases under review appears to matter (Davis, Haire, & Songer, 1993). Related work suggests that even if final decisions do not vary, having a woman on a court has nuanced effects (O'Connor & Segal, 1990).

Women in Executive Office

Since there has yet to be a woman president, a central stream of research in this area has focused on why there has not yet been one. Work on this question suggests public discomfort with the idea of a woman in the White House is key. Women presidential candidates also must battle gender stereotypes regarding expertise in foreign affairs (Lawless, 2004). Media coverage also works against women presidential candidates by focusing on their exceptionalism as women.

Other work has focused on the gendering of the institution of the presidency. Although all political institutions are gendered, Duerst-Lahti (1997) argues that the structure of the presidency is more "masculinized" than any other branch of government. This structural focus on hierarchy and command and control functions makes the Oval Office less accessible to women. The language used to represent the prototypical presidency in public discourse plays to masculine gender stereotypes among voters.

Another stream of research trying to explain the lack of women in the Oval Office has focused on the traditional pipelines men use for the presidency and how accessible they are for potential women presidential candidates. The verdict on this question is mixed. Women have been appointed to the cabinet increasingly over the last 40 years. However, the vast majority of women cabinet appointments have been in outer cabinet positions or further from the Oval Office. The result is that if and when they contemplate running for president, service as a member of a presidential cabinet is not as helpful as it would be for a man.

Women presidential candidates also face particularly unique challenges within the party system. The parties' interest in masculine candidates means that the more feminine the woman candidate, the harder it is for her to get the nomination (Conroy, 2007). Women presidential candidates also appear to face a different fund-raising viability standard.

Much less clear is whether women in the executive branch make a policy difference. When women are appointed to senior cabinet and statewide office positions, they are more likely to hire other women (Carroll, 1987). However, in terms of actual policy outcomes, women's potential effects on policy have been neutralized. They end up in traditionally feminized departments or as tokens with limited power. Research on leadership styles also suggests that gendered differences in leadership styles show up in executive differences, with women tending toward more

collaborative consensus building across stakeholders and men focusing on hierarchy. However, the limited number of women in statewide executive offices has made this question difficult to research.

Future Directions

Many possible new research directions arise from this review. Scholars should continue to identify whether various gender gaps will persist and whether current explanations will hold over time or whether new theories must be developed. Indeed, with regard to the gender gap and the small differences in political participation, it will be interesting to investigate whether these differences subside as women continue to achieve in both education and the workplace.

There is also a lack of research on candidate strategy relating to gender. For example, how do candidates strategically focus on their gender in their campaign communication through means like microtargeting and to what effect? It would also be useful to address more fully how gender and party stereotypes interact when the public evaluates female candidates. Additionally, does gender relate to strategies related to negative advertising?

Recruitment is another fertile area for future research. Given that qualified women view themselves as less qualified and the process to achieve office as more difficult, it is important to more carefully examine the process whereby qualified candidates in the pipeline become candidates. Scholars need to understand more systematically the role that political parties play in encouraging—or not—women candidates. Prior studies finding little relevance of gender on electoral outcomes or vote choice focus exclusively on general elections whereby female candidates have already been whittled down to a very elite, qualified group. Future research should work to understand the relevance of gender at various stages of recruitment prior to the general election.

There are several promising areas of research on Congress as well. One fruitful area for examination would be to look at whether the growing number of women in leadership positions (most notably, Speaker of the House Nancy Pelosi) results in women's increasing institutional effectiveness as Swers (2002) would suggest. The effects of the growing partisan polarization within Congress on women members should be examined as well. Intersectionality and the ways gender, race, and other forms of diversity work together within the halls of the U.S. Congress clearly also deserves future study. For example, have white congresswomen and African American congresswomen been influenced by institutional norms differently?

On the question of gender and the courts, there are growing numbers of women judges in both state and federal systems. This growing number offers the opportunity to more

fully evaluate gender and judicial decision making. For example, how do gender and race intersect within judicial decision making? Equally important, how does gender affect judges' ability to be appointed and then rise in the federal court system, the primary stepping stone to the Supreme Court? It would also be fruitful to examine whether and how gender stereotypes shape Senate judicial confirmation.

The most pressing question regarding gender and the presidency remains why no woman has yet been a major party candidate for president. The 2008 presidential election offers an intriguing story of almost that clearly requires further examination.

References and Further Readings

- Allen, D., & Wall, D. (1993). Role orientations and women state Supreme Court justices. *Judicature*, 7(3), 156–165.
- Bernstein, R. A., & Chadha, A. (2003). The effects of term limits on representation: Why so few women? In R. Farmer, J. D. Rausch Jr., & J. C. Green (Eds.), *The test of time: Coping with legislative term limits* (pp. 147–158). New York: Lexington Books.
- Bratton, K. (2005). Critical mass theory revisited: The behavior and success of token women in state legislatures. *Politics and Gender*, 1, 97–125.
- Burns, N., Schlozman, K. L., & Verba, S. (2001). *Private roots of public action: Gender, equality and political participation*. Cambridge, MA: Cambridge University Press.
- Burrell, B. (1994). *A woman's place is in the House*. Ann Arbor: University of Michigan Press.
- Burrell, B. C. (1993). Party decline, party transformation and gender politics: The USA. In J. Lovenduski & P. Norris (Eds.), *Gender and party politics* (pp. 291–308). Thousand Oaks, CA: Sage.
- Carroll, S. J. (1987). Women in state cabinets: Status and prospects. *Journal of State Government*, 60(5), 204–208.
- Caul, M. (1999). Women's representation in parliament. *Party Politics*, 5(1), 79–98.
- Conroy, M. (2007). Political parties: Advancing a masculine ideal in rethinking madam president. In L. C. Han & C. Heldman (Eds.), *Are we ready for a woman in the White House?* (pp. 133–146). Boulder, CO: Lynne Rienner.
- Darcy, R., Welch, S., & Clark, J. (1994). *Women, elections, and representation* (2nd ed.). Lincoln: University of Nebraska Press.
- Davis, S., Haire, S., & Songer, D. (1993). Voting behavior and gender in the U.S. court of appeals. *Judicature*, 77(3), 129–133.
- Dolan, K. (2001). Electoral context, issues, and voting for women in the 1990s. *Women & Politics*, 23(1), 21–36.
- Dolan, K. (2004). *Voting for women: How the public evaluates women candidates*. Boulder, CO: Westview Press.
- Dolan, K. (2008). Women voters, women candidates: Is there a gender gap in support for women candidates? In L. D. Whitaker (Ed.), *Voting the gender gap* (pp. 91–107). Urbana: University of Illinois.
- Duerst Lahti, G. (1997). Reconceiving theories of power: Consequences of masculinism in the White House. In M. Borelli & J. Martin (Eds.), *The other elites: Women, politics,*

- and power in the executive branch (pp. 11–32). Boulder, CO: Lynne Rienner.
- Freeman, J. (1986). The political culture of the Democratic and Republican parties. *Political Science Quarterly*, 101(3), 327–356.
- Gruhl, J., Spohn, C., & Welch, S. (1981). Women as policymakers: The case of trial judges. *American Journal of Political Science*, 25(2), 308–322.
- Han, L. C., & Heldman, C. (Eds.). (2007). *Rethinking madam president: Are we ready for a woman in the White House?* Boulder, CO: Lynne Rienner.
- High Pippert, A., & Comer, J. (1998). Female empowerment: The influence of women representing women. *Women & Politics*, 19(4), 53–67.
- Huddy, L., & Terkildsen, N. (1993). The consequences of gender stereotypes for women candidates at different levels and types of office. *Political Research Quarterly*, 46(3), 503–525.
- Kahn, K. F. (1996). *The political consequences of being a woman*. New York: Columbia University Press.
- Kathlene, L. (1994). Power and influence in state legislative policymaking: The interaction of gender and position in committee hearing debates. *American Political Science Review*, 88(3), 560–576.
- Kaufmann, K. M., & Petrocik, J. R. (1999). The changing politics of American men: Understanding the sources of the gender gap. *American Journal of Political Science*, 43(3), 864–887.
- Koch, J. W. (2000). Do citizens apply gender stereotypes to infer candidates' ideological orientations? *Journal of Politics*, 62(2), 414–429.
- Lawless, J. L. (2004). Women, war, and winning elections: Gender stereotyping in the post September 11th era. *Political Research Quarterly*, 53(3), 479–490.
- Lawless, J. L. (2009). Sexism and gender bias in election 2008: A more complex path for women in politics. *Politics and Gender*, 5(1), 70–80.
- Lawless, J. L., & Fox, R. (2005). *It takes a candidate: Why women don't run for office*. New York: Cambridge University Press.
- Mansbridge, J. (1986). *Why we lost the ERA*. Chicago: University of Chicago.
- Mansbridge, J. (1999). Should blacks represent blacks and women represent women? A contingent "yes." *Journal of Politics*, 61(3), 628–657.
- McDermott, M. L. (1997). Voting cues in low information elections: Candidate gender as a social information variable in contemporary United States elections. *American Journal of Political Science*, 41(1), 270–283.
- Mondak, J., & Anderson, M. (2004). The knowledge gap: A reexamination of gender based differences in political knowledge. *Journal of Politics*, 66(2), 492–512.
- Niven, D. (1998). *The missing majority: The recruitment of women as state legislative candidates*. Westport, CT: Greenwood Press.
- Norrander, B. (1997). The independence gap and the gender gap. *Public Opinion Quarterly* 61, 464–476.
- O'Connor, K., & Segal, J. (1990). Justice Sandra Day O'Connor and the Supreme Court's reaction to its first female member. *Women and Politics*, 10(2), 95–104.
- Palmer, B., & Simon, D. (2006). *Breaking the political glass ceiling: Women and congressional elections*. New York: Routledge.
- Paolino, P. (1995). Group salient issues and group representation: Support for women candidates in the 1992 Senate elections. *American Journal of Political Science*, 39(2), 294–313.
- Plutzer, E., & Zipp, J. F. (1996). Identity politics, partisanship, and voting for women candidates. *Public Opinion Quarterly*, 60(1), 30–57.
- Rosenthal, C. S. (1995). The role of gender in descriptive representation. *Political Research Quarterly*, 48(3), 599–611.
- Rosenthal, C. S. (2000). Gender styles in state legislative committees: Raising their voices in resolving conflicts. *Women and Politics*, 21(2), 21–45.
- Saint Germain, M. (1989). Does their gender make a difference? The impact of women on public policy in the Arizona legislature. *Social Science Quarterly*, 70, 956–968.
- Sanbonmatsu, K. (2002). *Democrats, Republicans, and the politics of women's place*. Ann Arbor: University of Michigan Press.
- Sanbonmatsu, K. (2006). *Where women run: Gender and party in the American states*. Ann Arbor: University of Michigan Press.
- Sapiro, V. (1981). Research frontier essay: When are interests interesting? The problem of political representation of women. *American Political Science Review*, 75(3), 701–716.
- Swers, M. L. (2002). *The difference women make: The policy impact of women in Congress*. Chicago: University of Chicago Press.
- Thomas, S. (1994). *How women legislate*. New York: Oxford University Press.
- Vega, A., & Firestone, J. M. (1995). The effects of gender on congressional behavior and the substantive representation of women. *Legislative Studies Quarterly*, 20(2), 213–222.
- Williams, M. (2007). Women's representation on state trial and appellate courts. *Social Science Quarterly*, 88(5), 1194–1204.

98

RELIGION AND POLITICS IN AMERICA

BRYAN A. S. RASMUSSEN

Collin College

One could easily make the argument that too much attention is paid to the United States by political scientists. If so, it is not as bad as it used to be. There was a clear ethnocentrism to early comparative political science. The effort apparent in much of that normative research was to try and build up what was great about American politics and see what the other nations had to do to be more like “us.” Political scientists today take a more empirical approach to studying the world. Still, a disproportional amount of attention continues to be paid to the United States. This may just be a function of the availability of data in the United States for political scientists to analyze. If this is the case, then recent attempts to expand the collection of public opinion and other data from around the world should help level the field. Regardless, it is easy to say that in many ways, a very stable political system in the United States has been overanalyzed, and many more dynamic political systems around the world have been significantly understudied.

In the field of religion and politics, the extra attention paid to the United States is definitely warranted. The United States is the outlier, the exception to the rule in many ways. It is a modern, rich, and developed nation that is remarkably religious. Why is the United States unique in this regard? This chapter reviews the significant literature produced over the past century attempting to answer this complex question. First, this chapter provides evidence that the United States is indeed unique and shows why this

fact is important. Then the chapter works to understand some of the specific characteristics of religion and politics in America that attract researchers to this field.

Evidence the United States Is Different

Are Americans really different than Europeans? Is the United States a significantly more religious country than most of the modern nations of the world? The United States certainly does not present itself to the rest of the world as a religious nation. A proper way to think about a nation’s core values might be to consider the reasons the people would be willing to go to war. Americans have shown a willingness to fight for democracy, capitalism, human rights, and other secular concepts. Would Americans go to war for Christianity? As of the writing of this volume, the United States is fighting wars in Afghanistan and Iraq. To some Americans, the purpose of these conflicts seems unclear, while some see Americans fighting for democracy or other noble cause. However, the consensus is clearly that Americans are not fighting on behalf of Christianity against Islam. These are not religious wars to Americans.

In addition, many Americans cannot conceive of the United States fighting a religious war. The Islamic extremists Americans are fighting do see it as a religious war, however. Interestingly, they do not see it as a war against

Christians, but a war against infidels, or secularists, that spread capitalism and threaten their traditional religious values. From this perspective, it would seem that the United States is perceived by much of the world as a secular nation.

The irony here is that by all measures, the United States is a relatively religious nation. From a macroerspective, the importance of religion on politics in the United States is easy to see. Public debates over social issues take on a much more religious tone in the United States than they do in Canada, England, or many of the other modern nations. Religion provides the foundation for mainstream arguments on political issues such as abortion, gay marriage, the death penalty, and many others (Brewer & Stonecash, 2007). Religion plays a significant role in electoral politics, especially in recent decades (Layman, 2001). The importance of religion in American politics may be intuitively clear, but the political scientist is not satisfied with mere observance. Political scientists must provide evidence to support their intuition. Only then can they begin to explore the causes and implications of a phenomenon.

The first step in measuring the religiosity of the American culture is to measure the religiosity of the American people from a microperspective. Traditionally, the most common measure of religiosity has been weekly church attendance. Although this number has been declining slightly in recent decades, still between 35% and 40% of Americans attend church at least once a week. This compares to an average of about 5% to 10% in many European countries. This statistic definitely supports the observation about Americans being religious, but caution is advised in reading too much into this number.

There are two main issues with this statistic: Is it accurate, and what does it mean? In terms of accuracy, there is evidence that church attendance numbers in the United States tend to be inflated (Presser & Chaves, 2007). The church has motivation to inflate attendance numbers, and the congregants may say they go more frequently than they really do because they want to look good. The causes for this motivation will become clear later, but evidence also suggests this motivation to inflate numbers does not exist to the same degree in other developed countries (Jelen & Wilcox, 1998). As a result, the difference in church attendance figures between these other nations and the United States may be overstated.

The other issue is what this statistic represents. It is definitely a very convenient measure. It has commonly been used in election studies and values surveys in the United States and around the world for decades, but does it illuminate what scholars want to know? At the basic level, the assumption that a religious person would go to church regularly and that the irreligious person would not go to church seems logical. However, this issue is more complex.

Does a person go to church because of a deep spiritual yearning, or is it a culturally accepted place for social

congregation? Is this person truly engaged in the organized religion of his or her choice, or are people just so-called pew potatoes who attend services out of habit or a sense of responsibility but are not engaged beyond that? Political scientists are realizing that the church attendance variable may be a good place to start, but it clearly has limits (Green, 2007). These days, it is often combined with variables measuring a person's beliefs (such as belief in God, heaven, or hell) or his or her willingness to attend church meetings outside of services to get a more complete picture of an individual's religiosity. Still, even with these other variables included, the United States still ranks among the most religious of the developed nations. Although belief in God has declined slightly (like church attendance) in the United States in recent decades, the number is still much higher than in much of Europe. No matter how one analyzes it, it seems religiosity is higher in the United States, setting it apart from most of the industrialized world.

Why Does It Matter?

To understand the current state of the literature in the field of religion and American politics, it is helpful to begin with the influence religion can have on individuals. It goes without saying that religion has always been important to people and has the capacity to influence how they live their lives. For many, religion is a cornerstone of who they are and plays a big part in many of the decisions they make. There is evidence that religion has been losing influence over the past 400 years and that this pattern continues today, although the extent to which this fact is true is up for debate. In any case, religion has always had, and continues to have, a very personal connection to many people and can therefore be used in many ways as a powerful motivator (or manipulator).

The fundamental importance of religion to an individual can have many manifestations in a society. One way to study it is how religion affects the institutions of the government. This institutional approach can focus on the structure of the institutions (whether a nation has a church-state or not) or how the politicians incorporate religion into the way they represent their constituents in a democracy. Another way is to examine the political behavior of the citizens, or the political culture of the nation, as represented by public opinion. In a pluralist nation like the United States, which has a strong civil society, religion can play a significant role in the exchange of ideas.

Religion has a unique ability to promote civil society and social capital (Smidt, 2003) and use these things to its advantage. First, religion has a deep meaning to many people. Therefore, a message delivered by a religious leader or given in a religious context can have a strong impact on individuals, causing them to change their behavior. Religions, to varying degrees, promote evangelism, or the

idea of sharing the good news with others. In addition, organized religions have a formal structure in place for regular meetings and the exchange of ideas. These factors combine to put religion in a unique position to influence the public.

Religion's role as a catalyst for social change is as old as religion itself. The power of the exchange of ideas that drove the Protestant reformation in the 15th and 16th centuries would soon move beyond religion and inspire the Enlightenment. This movement of liberal philosophers spanned the 16th and 17th centuries and represents an important turning point in our history. Although John Locke and many of these thinkers still wrote in a religious context, to not alienate their potential readers, they conceptualized a more secular world. Locke and Montesquieu envisioned a world where commerce was central to people's lives and secular education was to be of vital importance. They felt people had the ability to reason; therefore, they should be given freedom to run their own governments (via democracy) and invest themselves in science, commerce, or both, as they see fit. These philosophies, first incorporated by the founders of the United States and since throughout most of the world, have led to industrial revolutions and centuries of unprecedented advancement in science and technology. However, what does all this mean about the role of religion?

Religion had begun to lose its absolute authority in many societies. In addition, science and technology were beginning to provide answers to questions that could previously be answered only by religion. Friedrich Nietzsche, a German philosopher critical of the Enlightenment, went as far as saying that he envisioned a world where God would be dead. Science may not answer all the questions, but if commerce is there to keep people preoccupied, then they may lose their drive to delve into deep issues where religion could still provide guidance. The importance of religion to people would eventually be marginalized.

The dominant paradigm with regard to religion and politics through most of the 20th century centered around the idea that religion was headed toward extinction as a political force. Eventually, this idea was formalized into the secularization theory. Peter L. Berger (1999), one of the leading proponents of the theory, said that reason, scientific development, and bureaucratic specialization were among the factors that would eventually destroy religion as a political influence. Since this grand theory was the dominant paradigm for so long, not much was written about religion and politics for decades. Western Europe was becoming less religious by the decade, and the pattern of secularization worldwide, at least in more developed nations, seemed to confirm Berger's thesis.

This pattern never took hold in the United States. Religion continues to have a significant influence on American culture (Wald, 2003) and on how American institutions operate (Oldmixon, 2005). The United States has arguably led the world in industrialization, modernization,

and advancements in science and technology. Locke's idea of commerce becoming a central focus and taking people's attention off of deeper issues is as true in the United States as it is anywhere else. Despite all this, Americans are still as religious as much of the less developed world. How could this be?

What Makes Americans Different?

This is one of the most explored questions in political science literature in recent decades. For a long time, since the secularization theory dominated the literature, researchers felt the United States would eventually secularize also. Americans were just running behind for some reason. Events in American politics between 1980 and the present, with religion reemerging as a political force, obviously show this to not be the case. This leads to political scientists wanting to know why.

A number of hypotheses have been submitted and tested. One possible explanation is that the secularization theory is still accurate; it just needs a revision to explain the United States. Norris and Inglehart (2004) have a theory as to why the United States is still as religious as much of the third world that has yet to modernize and replace religion as the focal point of their society. Basically, the authors apply Inglehart's postmaterialist theory to the role of religion in the world. They posit that developed nations become more secular as they get more secure. As a result, postindustrial countries have less religious alignment and lower church attendance. The other part of it is that birth rates are lower in these countries; therefore, they see religion actually growing in the world because population rates are increasing faster in the non-postmaterial world, and this intensifies the line of conflict between the secure world and the third world. They try to explain the anomaly that is the United States by saying that the American minimal social welfare system fosters enough economic insecurity in cities and rural areas to cause religiosity similar to that in areas of the third world. This explanation seems problematic because one could easily argue that the homeless person living in a shelter in Los Angeles has a more comfortable life than half of the population of a country like Bangladesh.

One of the best developed theories in the literature concerning the religiosity in the United States is a supply-side model. Although many political scientists have contributed to this theory (Iannaccone, 1990; Jelen, 2000, among others), the most complete work is *The Churching of America, 1776–1990* by Finke and Stark (1992). It is a supply-side or market-driven model because the theory states that the religiosity in American culture does not come from an inherent demand for religion among the American people; it comes from the competition among the diverse suppliers of religion in the open marketplace.

The focus of this theory is the establishment clause of the First Amendment to the Constitution. At the most fundamental level, this clause guarantees a free marketplace for religious ideas. As a result, churches publicly compete for members, and this overt competition creates a religious fervor in the community. Ironically, in many cases today, churches do not compete for membership using differences in interpretation of scripture or other doctrinal differences. Instead, the competition takes a more secular form, as the churches advertise location, service times, opportunities for children, or the quality of their choir or other programs. This competition brings people in the doors, and then the messages they hear reinforce the religiosity in the culture.

An interesting side effect of this competition is the relatively high percentage of self-proclaimed atheists or irreligious people who live in the United States. Just as the free market allows individuals to be whatever religion they want, it also allows them to choose not to be any at all. Depending on the poll, usually between 15% and 20% of Americans claim to be atheist, among the highest rates in the world. So although the market model has been shown to promote religiosity, it can also work to promote an aversion to organized religion.

The other manifestation of this phenomenon to note is that many of the fastest growing churches in the United States today are nondenominational churches. These people are religious, but they are averse to the traditional denominations, so they have stepped out of the box in a different way that the market allows. This phenomenon in American culture seems to go beyond religion as well. Although Americans used to take pride in certain labels, they now seem to take pride in not being labeled. An example of this outside of religion would be partisanship. Polls have shown a decline in partisanship among Americans over the past 40 years. Political scientists are studying, however, whether this reflects actual changes in people's voting behavior or if it is just that Americans are less willing to claim to be a Republican or a Democrat although their voting behavior remains distinctly partisan.

Finke and Stark (1992) support the market model by reviewing the history of religion in America. It is important to note that for most of American history, the religion was Christianity. It was not until the mid- to late 20th century that significant minorities of other religions and irreligious people began to form. Finke and Stark show that at times in American history when one sect of Christianity grew stronger, resulting in a decrease in competition, overall religiosity in the culture decreased. However, when religious competition increased, so did religiosity.

There are many reasons for this. During times of greater competition, churches are willing to try different things and be more aggressive pursuing congregants. There is excitement in the air as different denominations vie for membership. When one sect gets the upper hand, the leaders of that group tend to get comfortable with their

position. They put less energy into bringing people in and more energy into managing the people they already have. The energy around religion decreases, and the smaller sects, at a significant disadvantage, feel less able to try and compete.

Finke and Stark (1992) demonstrate that churches that have generated the most religious energy throughout our history were, ironically, the churches that expected the most commitment and energy from their constituents. As a general rule, when groups form and try to expand, the natural tendency is for them to make engagement as easy as possible to lure more people in. This example shows the error in that thinking. Many of the rapidly growing religious groups were new denominations. As they struggled to get started, the leaders and the followers had to make a deep commitment to get it to work. In other cases, the denomination's approach to religion required a stronger commitment from the congregant. An example of both of these would be the Methodists. Methodism developed in the mid-19th century as circuit riders traveled from community to community evangelizing and preaching the Word. It was a very energetic approach to religion that required a lot from the lay leaders when the minister was not there. Over time, the Methodists have developed into a mainstream Protestant denomination and, arguably, have lost some of their energy. They, like other mainline Protestant denominations, have actually seen their membership shrink in recent decades while groups that require a greater personal commitment, such as Southern Baptists and Mormons, continue to grow.

The strength of this market model is not limited to Christianity or the United States. Islamic nations around the world tend to form church-states. However, the growth of Islam in recent decades in the United States is another example that if an active commitment is required, a religion can do well in a free-market system. Attempts to apply the market model to other nations in Europe have met with limited success (Jelen & Wilcox, 1998), while recent attempts to apply this model to Latin America have proven more successful (Gill, 1998). Although the market model seems to fit the United States very well, scholars today are studying the theory's generalizability.

There are a number of aspects of this market theory worth exploring. First, it is important to recognize that this is a top-down, elite-driven theory. Again, religiosity is driven by the church leaders, not the people. Although this is a solid theoretical assumption, and one that often fits the facts, it clearly is not always true. As long as a country has a strong civil society, there will always be opportunities for issues to rise from the bottom as well as trickle down from the top. Americans are a communal people, willing to develop a strong civil society and to allow religion to play a role in that process. The fact that this is true was first observed by Alexis de Tocqueville (1966) in the 1840s and continues to be observed today. Although some see signs that civil society is breaking

down in America today (Putnam, 2000), the tradition remains a strong part of the culture.

Second, using this top-down approach allows political scientists to use a rational choice model to explain the behavior of churches. This was a difficult adjustment for political scientists. For a long time, the idea of combining rational choice and religion was considered inappropriate, to say the least. Religious people, by their very nature, are assumed to be not rational, since their behavior is driven by a belief in something mystical. Over the years, political scientists have begun to realize that religious people do act rationally and that religions in an open market can definitely be seen as acting rationally. Gill's (1998) study of the ability of evangelical Protestants to gain ground in Catholic Latin America is a strong example of this use of the rational actor model.

Another aspect of this theory that has garnered a great deal of research over the past decade is the ability of the church to act as an agent of socialization and manipulate people's views. This is fundamental to the market theory. The competition may get people into the churches, but that only matters if the church is then able to have an influence on its congregation and make them more religious in their worldview and, potentially, in their politics. Scholars have taken a number of approaches to this question.

The first step is to look at the clergy. In some religions, the opinions of the individual giving the message matter less than in other religions. Catholicism, for example, has traditionally been a more centrally run church, with the Word coming from God, through the pope, to the local priest. This would not give the local priest much opportunity to deviate from Catholic doctrine. The fact is that even in Catholicism, scholars have seen a split in recent decades between more orthodox Catholics and less orthodox Catholics. Approximately half of American Catholics support the Democratic Party, which has a prochoice platform, while the Vatican clearly takes a prolife stance.

If there are opportunities for deviation in the Catholic Church for local priests and congregations, there are certainly opportunities for Protestant clergy. Protestants lessen the role of the church and focus more on the individual's relationship with God. Therefore, there is much more room for individual interpretation by local clergy and congregations, both among Protestant denominations and within a denomination. If the split within a denomination is severe enough, members can form another denomination, as the Lutherans are currently doing over the gay clergy issue.

One thing this chapter has already established is that many of the fundamental differences among religions and denominations do not seem to be the focus of sermons one might hear on Sunday, because of the market system. One is unlikely to hear a sermon from a Methodist pastor outlining the doctrinal difference between them and the Lutherans, for example. From this perspective, the politics of the individual clergy matters, since they have some leeway

in how they deliver the message. Guth, Green, Smidt, Kellstedt, and Poloma (1997) and Smidt (2004) are two examples among many of political scientists recognizing the importance of the politics of the individual clergy. Both recognize patterns showing clergy of the different denominations having different political views. Political views of clergy are important, and there are definite patterns in partisanship and other things to be recognized among Protestant clergy in particular.

The next question concerns clergies' ability to pass their views to their congregations. They do have a certain authority and influence over many people as men or women of God. At the same time, however, they are limited in their ability to engage in politic issues for two main reasons. The first reason is their desire to maintain tax-exempt status. For a church to be tax exempt, they cannot engage in politics in any direct way. They must remain neutral. It seems like every election year, there are churches in different communities trying to test the boundaries of these laws, but the limitation is considerable.

The other main limitation is their desire not to offend potential congregants. In an effort to grow, many mainstream churches need to be willing to expand their acceptance of diversity of opinion. They cannot be too specific in support of a party or an issue if they want to appeal to a broader audience. This limits a church's desire or ability to take a strong stand on any social or political issues. Whether trying for the broadest possible appeal is a good idea for a church is a debate for another time. As a result of this goal, however, people who attend church regularly will answer collectively about their belief in God or the need to love their neighbor (whether they actually do this is another subject), but the ability of the clergy to socialize on specific social issues is very limited.

Another part of the issue with clergy getting political is the fact that the core messages sent from churches do not line up well with the American party system and political ideologies. For example, churches are, by their nature, communal organizations that promote collective responsibility. A church might feel it needs members to work together to help the less fortunate. This sense of collective responsibility is a more liberal approach to problem solving as opposed to a more conservative approach focused on individual responsibility. At the same time, many churches promote more socially conservative issues, with stands against abortion, gay marriage, euthanasia, and others. If the church is sending messages that do not easily align with a certain political ideology or political party, the congregant is left to rationalize what message they choose to receive, allowing for a level of selective perception.

The other limitation is the extent to which the congregant is engaged enough to receive certain messages at church and have these messages change their opinions on a subject. Many political scientists continue to explore this issue (Layman, 2001; Wilson, 2007). This issue ties back into the previously explored idea about the level of expectation

the religion has for its members. Generally, the higher the expectation, the more the member is engaged and the more likely the message from the clergy will have a significant impact on the actions of the congregant.

This market model has many aspects and implications that political scientists continue to explore. The main cause for this open competition among religions is the establishment clause. Although the clause seems simple enough, one of the big issues political scientists continue to explore is the controversy over how it is to be interpreted.

Accommodationists Versus Separationists

The First Amendment of the U.S. Constitution (1787) contains a freedom of religion clause (the free exercise clause) and a freedom from religion clause. The free exercise clause does not generate much controversy. The fact that people can be whatever religion they want to be, or no religion at all, is almost universally accepted. Even in nations around the world that have church-states, many of them have freedom of religion clauses in their constitutions (Americans may wonder about the legitimacy of those clauses, but it works for them).

The only controversies that come up here usually concern one of two questions. First, what constitutes a religion? If religions qualify for tax-exempt status and other benefits, there needs to be some standard for this. Certainly Christianity, Judaism, Islam, Buddhism, Hinduism, and all the main world religions would qualify. However, one can think of an example of a group that some consider to be a religion and others do not. What qualifies as a religion?

The other question that the courts address from time to time is this: What can be done in the name of a religion? Generally speaking, people cannot violate laws in the name of religious practice. This has been more controversial in the past, but an occasional issue will still come up today. The fact is, however, the free exercise clause does not generate much attention from political scientists.

The establishment clause is different. The clause states that "Congress shall make no law respecting an establishment of religion" (U.S. Constitution, 1787, First Amendment). The fact that this clause prohibits the United States from establishing a church-state like many other nations have is clear. The United States was founded in part by religious refugees. Granted, they represented minority Christian sects trying to escape majority Christian sects, but still they were outcasts. Banning a church-state was a very important part of religious freedom to them.

Still today, the idea of a church-state in the United States does not have any wide appeal. Around the Western world, many nations that have church-states are moving away from them. Besides, if the United States were to establish a state-church, which religion or denomination

would it be? It would certainly be Christian, given that almost 70% of Americans are Christian. However, they are split among multiple Protestant denominations and Catholics. Remember, the United States also has a relatively large percentage of atheists that have a significant say in the American pluralist political system.

In any case, this is not the controversy. The controversy centers around the extent to which church and state should be separated. This debate centers around two groups that political scientists have labeled accommodationists and separationists (Jelen, 2000).

Separationists interpret the establishment clause to mean there should be a complete separation of church and state in the United States, or as Thomas Jefferson wrote, "a wall of separation." Separationists fight for no prayer in schools and for keeping the Ten Commandments out of public buildings. In recent years, a separationist appellate court in California decided that the phrase "under God" in the Pledge of Allegiance should be taken out if kids are asked to say the pledge in school. This decision was set aside and not enforced, but it provides a good example of the goals of separationists.

Irreligious people are naturally going to be separationist. If one does not like religion, why would one want religion to be engaged in some way with government? On the other side, however, there are many religious people who are separationists. The reason for this is that they believe religion and government to be fundamentally incompatible. After all, religion is about right and wrong, and democratic politics is about compromise. These separationists fear that government will corrupt their religion or at the very least water down its message. They find examples in Europe, where the church-state governments there take relatively liberal positions on many social issues.

The American law of the land is currently separationist. The current court precedent, *Lemon vs. Kurtzman* (1971), says that the government's action must have a secular legislative purpose, it must not have the primary effect of either advancing or inhibiting religion, and it must not excessively entangle government with religion. This precedent has been narrowed and defined by the courts over the years, but it remains the legal foundation for all establishment-clause issues.

Accommodationists, on the other hand, want government's approach to religion to be one of so-called positive neutrality (Jelen, 2000). The idea is not to pick one religion over another but to recognize that general religious values can be used to support government and provide a stronger foundation for government action. An example of an accommodationist would be former president George W. Bush. He is a very religious man who was not shy about using his faith to guide his decisions about the so-called right thing for government to do. He also pursued a number of faith-based initiatives, including allowing students to go to religious schools with government money and allowing

churches to distribute government social welfare money in local communities. The idea is that religion can provide a deeper meaning for what government does and more motivation for citizen engagement.

There is a fundamental aspect of this debate that cannot be ignored. The accommodationist position implies that it is acceptable for a religious people, Christians specifically, to engage their faith in politics. Mohammed was a spiritual leader, a military leader, and a political leader, leading to the belief that the Islamic state is a natural fit for Muslims. Jesus was a spiritual leader, who made a point of saying his followers should respect God and the state and keep them separate. As such, many Christians have been reluctant to engage their faith in politics.

The Rise of the Evangelicals

In the past, studies of religion and party politics in the United States focused on the differences in the religious sects. Traditionally, Protestants were viewed as more likely to be Republican, and Catholics and Jews as more likely to be Democrats (Green, 2007, calls this the “old gap”). A number of historical, economic, and social theories were put forth to explain this division, focused largely on the New Deal Era in American history (Layman, 2001). With the partisan realignment surrounding civil rights beginning in the 1960s and the emergence of the Christian Right in the 1980s, there is a debate as to whether this traditional approach is still appropriate (Green, 2007).

Increasingly, scholars are taking the view that what religion a person subscribes to is less important in determining partisanship than how religious they are. More and more, evidence shows that more religious people, regardless of religion, tend to support Republicans, and less religious people (and irreligious people) tend to support Democratic candidates (Green, 2007, calls this the new gap). This new gap became a significant force in presidential elections beginning in 1992 and has remained consistent or grown bigger since then, causing many to see it as a possible partisan realignment (Fiorina, Abrams, & Pope, 2006).

There are a number of possible reasons for the alignment of more religious people with the Republican Party. It may be as simple as a shift of the Republican Party platform over the past 25 years toward Christian positions on a variety of issues. To say that Republicans better represent religious interests and Democrats better represent secular interests may be generally accurate at this time, but it does not tell much of the story. It also does not explain how or why this convergence between religious values and Republican issues has taken place.

One way to explain the convergence of traditional religious values and the Republican Party in recent decades is as an elite-driven rational choice model. The theory states that the religious elites found an opportunity to expand their influence by entering politics through President

Reagan and the Republican Party in the 1980s (Layman, 2001). This also gave Reagan an easily identifiable base that could be targeted and mobilized for his reelection bid in 1984 (Holbrook & McClurg, 2005). What began as a marriage of convenience for ambitious elites continued to progress as Republican positions began to further align themselves with Christian values, and Democratic positions, in response, tended toward secular positions.

Interestingly, the Christian Right entered politics in the 1980s very aggressively talking in religious terms about what people needed to do and what would happen to the people if they did not do it. They struggled with this approach in a culture that sees itself as secular. As a result, they modified their approach and began speaking of things like family values. Although this has proven to be more effective for them politically, it has created a dynamic where the Christian majority in our country feels it has the same freedom of speech discrimination issues that might be encountered by an oppressed minority (Wald, 2003).

Some have gone so far as to suggest Americans have a culture war in their country as religious people align with the Republican Party against secular Democrats over cultural issues (Brewer & Stonecash, 2007). Others say this culture war is a myth perpetuated by the political elite designed to motivate people to participate in elections (Fiorina et al., 2006). In any case, the role of religion in politics has definitely increased since Ronald Reagan ran for president in 1980.

Future Directions

The field of religion in American politics is very dynamic and exciting. Most of the research has been done within the past 30 years, and there is a lot more to do. The theories outlined in this chapter are by no means complete, nor do scholars know their full implications.

The possibilities for future research are limitless. After September 11, 2001, the recognition of the importance of religion in politics continues to grow. There are two areas this section points to specifically, however, as avenues for continued research.

The first is the effect of Federalism on religion and politics. The state is an important unit of analysis for electoral purposes. After all, American elections take place at the state, district, or local level. Americans have only one national election, and even the presidential election is decided on a state-by-state basis because of the Electoral College. Therefore, states are very important units of analysis for voting behavior.

However, is there something distinct about a state that would lead it to have a unique red or blue culture? This is a federalism argument, and there is extensive debate on this issue in the literature. In arguing for state importance, a core concept is that all politics are local. It is argued that local and state politics drive the federal system and that the

national government is at its core a sum of the state governments (Chibber & Kollman, 2004). Most of the laws that affect Americans' daily lives are made at the state and local level, and there is considerable variation among the states in laws regarding important social and economic issues. From this perspective, the states could easily develop independent red or blue identities.

The other side of the argument is that our states have been overrun by the federal government over the past 200 years and have little if any real power left. In this case, maybe the variation in culture seen across the country is not on a state-by-state basis but better defined in a regional context as Green (2007) does. In many cases, the neighboring states share a similar history and appealed to certain groups of immigrants and business interests that helped define a regional culture. It is easy to recognize that the United States is made up of a number of subcultures. For example, it is easy to see a general cultural difference between the Northeast and the South. The question remains: Does one define these differences in terms of states or regions?

The Federalism issue is one political scientists have been debating, and there is a long way to go. The other issue that is ripe for future research has been recently framed by Peter Berger and Anton Zijderveld (2009). The same man that drove the secularization theory and contributed to its demise has recently given a new framework to consider religion and politics. *In Praise of Doubt* (Berger & Zijderveld, 2009) is a theoretical book where the authors discuss the relativism that is perceived as fundamental to modernity. The authors point out that the absolutism that is often applied as a criticism of religion can also be found in relativism. Therefore, relativism should not get credit for modernization, but the key instead is the mere capacity to doubt. Although the existence of doubt may change religions' role in a culture, it does not eliminate religion as a significant cultural force. This may provide a new paradigm with which to look at the role of religion in politics. The bottom line is this: Although scholars have made some progress understanding the affect of religion on politics, there is still much work to be done as they continue to develop and test new theories.

Conclusion

In many ways, the study of religion and American politics is relatively new. The market model goes a long way toward showing why the United States is not cooperating with the secularization theory, but there is much work to do. In addition, the dynamic political scientists are endeavoring to study is constantly changing. From the rise of the Evangelicals in the 1980s to September 11, 2001, events continue to challenge their understanding of the relationship. Once they begin to conceive of a party realignment around religion (the Republicans centering around more

religious people and the Democrats centering around a more secular approach), then comes the election of 2008 where the Republicans nominated a candidate unappealing to religious conservatives and the Democratic candidates were openly discussing their faith on the campaign trail. Political scientists need to try and understand the basic theories in the field as quickly as possible before they change again.

References and Further Readings

- Berger, P. L. (Ed.). (1999). *The desecularization of the world: Resurgent religion and world politics*. Washington, DC: Ethics and Public Policy Center.
- Berger, P. L., & Zijderveld, A. (2009). *In praise of doubt: How to have convictions without becoming a fanatic*. New York: HarperCollins.
- Brewer, M. D., & Stonecash, J. M. (2007). *Split: Class and cultural divides in American politics*. Washington, DC: CQ Press.
- Chibber, P., & Kollman, K. (2004). *The formation of national party systems*. Princeton, NJ: Princeton University Press.
- Finke, R., & Stark, R. (1992). *The churching of America 1776-1990*. New Brunswick, NJ: Rutgers University Press.
- Fiorina, M., & Abrams, S. L., & Pope, J. C. (2006). *Culture war? The myth of a polarized America*. New York: Longman.
- Gill, A. (1998). *Rendering unto Caesar: The Catholic Church and the state in Latin America*. Chicago: University of Chicago Press.
- Green, J. C. (2007). *The faith factor: How religion influences American elections*. Westport, CT: Praeger.
- Guth, J. L., Green, J. C., Smidt, C. E., Kellstedt, L. A., & Poloma, M. M. (1997). *The bully pulpit: The politics of Protestant clergy*. Lawrence: University of Kansas.
- Holbrook, T. M., & McClurg, S. D. (2005). The mobilization of core supporters: Campaigns, turnout and electoral composition in United States presidential elections. *American Journal of Political Science*, 49, 689-703.
- Horowitz, D. L. (2000). *Ethnic groups in conflict*. Berkeley: University of California Press.
- Huntington, S. P. (1996). *The clash of civilizations and the remaking of world order*. New York: Touchstone.
- Iannaccone, L. (1990). Religious practice: A human capital approach. *Journal for the Scientific Study of Religion*, 29, 297-314.
- Jelen, T. G. (2000). *To serve God and mammon: Church state relations in American politics*. Boulder, CO: Westview.
- Jelen, T. G., & Wilcox, C. (1998). Context and conscience: The Catholic Church as an agent of political socialization in western Europe. *Journal for the Scientific Study of Religion*, 37, 28-40.
- Jelen, T. G., & Wilcox, C. (Eds.). (2002). *Religion and politics in comparative perspective: The one, the few and the many*. Cambridge, UK: Cambridge University Press.
- Layman, G. (2001). *The great divide: Religious and cultural conflict in American party politics*. New York: Columbia University Press.
- Levine, D. H. (1986). Religion and politics in comparative and historical perspective. *Comparative Politics*, 19, 95-122.
- Lijphart, A. (1999). *Patterns of democracy*. New Haven, CT: Yale University Press.

- Mitchell, J. (2007). Religion is not a preference. *Journal of Politics*, 69, 351–362.
- Mitchell, J. (2008). A reply to my critics. *Journal of Politics*, 70, 880–883.
- Norris, P., & Inglehart, R. (2004). *Sacred and secular: Religion and politics worldwide*. Cambridge, UK: Cambridge University Press.
- Oldmixon, E. A. (2005). *Uncompromising positions: God, sex and the U.S. House of Representatives*. Washington, DC: Georgetown University Press.
- Pals, D. L. (2006). *Eight theories of religion*. New York: Oxford University Press.
- Presser, S., & Chaves, M. (2007). Is religious service attendance declining? *Journal for the Scientific Study of Religion*, 46, 417–423.
- Putnam, R. (1994). *Making democracy work*. Princeton, NJ: Princeton University Press.
- Putnam, R. (2000). *Bowling alone*. New York: Simon & Schuster.
- Smidt, C. (Ed.). (2003). *Religion as social capital*. Waco, TX: Baylor University Press.
- Smidt, C. (Ed.). (2004). *Pulpit and politics: Clergy in American politics at the advent of the millennium*. Waco, TX: Baylor University Press.
- Tocqueville, A. de. (1966). *Democracy in America*. New York: HarperCollins.
- U.S. Constitution. (1787). Available at <http://www.usconstitution.net/const.html>
- Wald, K. D. (2003). *Religion and politics in the United States*. Lanham, MD: Rowman & Littlefield.
- Weber, M. (1930). *The Protestant ethic and the spirit of capitalism*. London: HarperCollins.
- Wilcox, C., Wald, K. D., & Jelen, T. G. (2008). Religious preference and social science: A second look. *Journal of Politics*, 70, 874–879.
- Wilson, J. M. (Ed.). (2007). *From pews to polling places: Faith and politics in the American religious mosaic*. Washington, DC: Georgetown University Press.

99

LGBT ISSUES AND THE QUEER APPROACH

ROBERT KULPA

Birkbeck College

Intimate desires and erotic feelings seem to belong so obviously to the domain of the private sphere of each human that recognizing sexuality as an important element of politics—that is, of public activity—has not been an untroubled process in American social, political, and economic histories. Indeed, sexuality continues to be one of the most contested issues in modern politics, so much so that it is almost impossible to talk about modern liberal democracy, with all its ideological baggage, claims to human rights, and individual freedoms and liberties, without addressing the issue of sexuality. Without doubt, one of the factors that enabled this shift was the rise of identity politics in the post–World War II world. As such, the rise of lesbian, gay, bisexual and trans (LGBT) identity politics is part of wider structural shift in political organization of contemporary democracies.

The primary focus of this chapter is to introduce the reader to the contemporary debates surrounding LGBT issues in U.S. politics. It is done by looking at LGBT movements and major issues raised by those communities. First, a historical context is offered, followed by the overview of major problems raised by LGBT movements. These are composed of education, representation, legal regulations, health, and strategies of political action. The next part, “Queer Approaches,” develops the conceptual side of this chapter, discussing identity politics, market economies, sexual citizenship, nationalism, terrorism, and neoimperialism. This chapter concludes with an overview

of its contents, indicating development of LGBT and queer (Q) politics and its futures and suggests further readings.

Histories of Lesbian and Gay Movements in the United States

The Postwar Period

The years from 1950 to 1969 can be best understood under the reformist flag of homophile activism and could be best characterized by the slogan from the era: “We are just like everyone else.” At the time, most of the sociological, psychological and medical, and legal studies saw the term *homosexuality* as a form of deviance or pathology. The political powers of the cold war period saw homosexuality as threatening to the state as communism (most notable here is U.S. Senator Joseph McCarthy’s so-called witch hunts, hauling people up before the House Committee on Un-American Activities between 1934 and 1975). But it was also the time when Alfred Kinsey published his reports on the sexual behavior of Americans (in 1948 and 1954), which showed what was actually happening in the bedrooms of Americans—and it was not as puritan as the official discourses about it suggested. It was also a time when the first gay (e.g., the Mattachine Society [MS]) and lesbian (Daughters of Bilitis [DoB]) organizations appeared. The actual word used by the

activists of the time was *homophile*, which reveals the major goals and attitudes of homosexual people of the time. The term was coined by activists to downplay the sexual aspect (homosexuality) of sexual identity, thus attempting to erase the difference and put more stress on the sameness with the heterosexual majority. Homophile organizations hoped for assimilation by using nonaggressive methods and accepting social norms (of gender, class, and race).

Moreover, not only did homophile organizations not challenge social norms, but they also acted in a way that strengthened them, most notably those concerning sex and gender roles. Thus, DoB and MS in their magazines advised readers to look and dress “properly,” reinforcing standards of masculine men and feminine women. This can be seen as one example of homophile groups focusing on education as a crucial strategy of their movement.

This aspect is also present in the adoption of an expert-system approach: attempting to depathologize homosexuality by convincing so-called experts (lawyers, doctors, and government officials) of the normality of homosexual people. Experts would then, by the token of their expertise, influence and change society.

Overall, the profile of the homophile movement of the 1950s and 1960s could be generally summarized as assimilationist, characterized by a rather tame approach (by today’s standards), involving the liberal pledge of sameness and acceptance of rigid gender expressions. Sexual identity was seen as totally confined to the private, personal bedroom life. Any attempt to make homosexuality a matter of public interest was done by separating it from any other aspect of identity and normalizing gender expressions according to dominant norms.

Gay Liberation and Lesbian Feminism

The opening up of another chapter in lesbian and gay politics is commonly ascribed to the riots around the Stonewall Inn drag and gay bar, in New York, in June 1969. This period is usually referred to as *gay and lesbian liberation* and characterized by such slogans as “Out of the closets and into the streets” or “Gay revolution now.”

The significant shift of this decade of lesbian and gay activism is a move away from the shame and pledge of similarity that dominated homophile politics and to stress instead difference, pride, and rage, which became characteristic of liberation movements of the time. Notably, the shift came on the wave of American counterculture, of which sexual revolution, second-wave feminism, black-power liberation, and antiwar movements are prime examples. Similar to those of other liberationist groups and movements of the time, lesbian and gay strategies were those of direct action: street protests, intervening in conferences and meetings, and disturbing other public gatherings. Clearly, the stress was on public activity; sexuality was celebrated and raised to serve as a tool of the personal and group liberation, as a vehicle for change.

Lesbian feminism grew out of disappointment with the androcentrism of gay liberation and heteronormativity (heterosexuality as a norm) of early second-wave feminism. It promoted women-only spaces as safe and secure enclaves for women to nourish and cherish their bonding, self-education, and consciousness-raising work. Some of the adopted strategies were those of separatism (although highly contested) and education, through magazines like *Furies*.

The flagship manifesto, *Women Identified Women*, written by the collective Radicalesbians in 1970, was also groundbreaking. It opened space for considering lesbianism and identification as a lesbian as a political act, not only a sexual desire. The idea was later reworked in another pivotal text, Adrienne Rich’s (1980) “Compulsory Heterosexuality and Lesbian Existence.” However, political lesbianism, as it later came to be known, was also criticized by some other feminists. They felt that taking sexuality away from lesbianism was sacrificing desire and eroticism for the sake of common (all-women) identity (as a means of liberation). The uneasy relationship between feminism and sexuality was evident again during the so-called 1980s sex wars, notably during the feminist debates over pornography, S&M practices, and the nature of sexual violence.

Rooted in the counterculture of the late 1960s and early 1970s, lesbian and gay liberation had a vision of heteronormative oppression (a term coined later) as an outcome of patriarchal society. This, in turn, was sustained by racism and Western (American) global imperialism. Such vision helped to establish lesbian and gay identity politics. It became a way of conceptualizing homosexual people as similar to ethnic minorities, thus adopting an essentializing (i.e., stressing the essence and nature and biological determinacy), ethniclike vision of homosexuality as a fixed identity.

Just as the word *homophile* was adopted by the activists of the 1950s and 1960s, so the 1970s brought the word *gay* as the preferred description. But while homophile movements wanted to repair social relations and present homosexuality as acceptable for the heterosexual majority, gay liberation emphasized and celebrated difference and rejected the concept of so-called normality. The ultimate aim was to eradicate gender roles and transform the family as a social institution, thereby ending homophobic (fear of homosexuality leading to abusive behavior) violence—and most significantly, reconceptualizing sexuality not in terms of reproduction and social status, but pleasure and relationship. This, in turn, was a crucial moment for establishing bisexuality as factor of politics (see subsequent paragraphs).

The move toward conceptualizing homosexuality as the unifying and predominant base of identity, on which a social movement could be built, can be described as strategic essentialism. It was introduced on a large scale at the beginning of the 1970s but took over and dominated lesbian and gay politics by the end of the decade.

The late 1970s also saw the shift from revolutionary lesbian and gay (cultural) politics into a more formal and

structural (political) type. It was reflected in the vanishing of the Gay Liberation Front, a loose, nonhierarchical, and not very formalized group or organization. At the same time, there appeared nongovernmental organizations like National Gay (and Lesbian—added later) Task Force, which were formalized, structured, and funded, with full-time employment.

Overall, lesbian and gay liberation movements brought the radicalization of politics, invasion into the social and cultural status quo, rejection of puritan visions of sexuality and gender, and a stress on difference. In the New Left spirit, it also stressed a need for coalitions with other marginalized and discriminated groups. It was also the time when a strong, identity-based model of activism established itself as the dominant one. By the end of the 1970s, along with some relative successes of previous struggles, came a new approach to gay rights, focusing on political rather than cultural issues. It was more of a single-issue type of politics and predominantly focused on gay rights as part of human rights. This gay rights approach dominated lesbian and gay politics in the 1990s, though not without controversy (see the subsequent section titled “Queer Approaches”).

AIDS

The shift within lesbian and gay politics was also the effect of the social backlash that came in the 1980s with the appearance of HIV and AIDS, on the one hand, and the political hegemony of New Right conservatism, winning power in the United States and the United Kingdom, on the other. However, in retrospect, it can be said that these factors were also catalysts for the emergence of the radicalized HIV and AIDS movement and later queer politics. These, however, developed a different set of approaches in sexual politics. The first half of the 1980s brought about crystallization of identity politics, emphasizing gay identity as the key factor in social mobilization. However, governmental nonresponsiveness to the HIV and AIDS epidemic soon made clear that the gay rights approach, putting its trust into state-sanctioned channels of lobbying and pressurizing, was not enough.

In the second half of the decade, new organizations such as the AIDS Coalition to Unleash Power (ACT UP) and Queer Nation emerged. This is often seen as the beginning of queer politics. These groups brought back more radical methods of social activism and stressed that identity-based politics were too rigid to cope with multi-level, all-encompassing social issues like HIV and AIDS.

In addition, the rise of Christian fundamentalism in the United States helped to spread the backlash politics of fear, creating moral panics. Christian fundamentalism presents itself as defending family values from homosexuality (seen as a threat to a traditional morality and society), opposing the special rights and special interest groups discourse that they saw lesbian and gay people were demanding and imposing on the straight majority. Christian fundamentalist

groups proved to be an especially vicious opponent of LGBT movements.

If, as the famous ACT UP slogan from that time was proclaiming, “Silence = Death,” then the antidote was to be as noisy as possible. Gay activism around this time employed strategies that brought attention to issues that Reagan’s and Bush’s governments were reluctant to address, and this type of high-profile activism became one of the core features of LGBT politics. ACT UP and Queer Nation, like the Gay Liberation Front before them, adopted loose, nonhierarchical structures and focused their activities on public space, such as kiss-ins (groups of activists gathering together in public spaces, like shopping malls, and embracing mass kissing), die-ins (staged dying to expose religious bigotry as poisonous), and video activism. Groups like Testing the Limits Collective that made the famous *Voices From the Front* video diary and documentary in 1992 were close to HIV and AIDS organizations, working in collaboration for the common cause.

The B and T in LGBT

The acronym LGBT points toward four gender and sexual identity categories: lesbian, gay, bisexual, and trans. They are usually treated as having similar interests and belonging to the same universe. However, the relationship between lesbian and gay communities and movements, on the one hand, and bisexual and trans, on the other, was neither easy nor painless.

The Bisexual Movement

The beginnings of the organized bisexual movement can be dated back to the late 1960s. However, it really began to develop by the end of the next decade, flourishing and firmly establishing itself on the political map of identity politics by the end of the 1980s. The rise of a bisexual movement is the result of various factors.

First, it derives from bisexual people’s disappointment with the lesbian and gay movements. Initially, bisexuals invested their energy, trust, and hopes in the gay movement; it seemed natural to align the two causes. However, the lesbian and gay movements soon proved to pay little attention to, or have little interest in, the specific needs of bisexual communities.

Moreover, bisexual people were often treated with suspicion and hostility, perceived as traitors of the sexual-liberation cause. Bisexual women found that sleeping with men was against lesbian feminist orthodoxy of that time. Similarly, bisexual men found that having sex with women was seen by gay men as denying homosexuality and lacking the courage to come out. This exclusion was later taken up by queer activists as an example of how identity politics may be constraining for those who actually want to use it for their liberation.

It is thus somewhat paradoxical that while lesbian and gay movements were fighting for their liberation from the heteronormative society, these oppressed groups did not avoid the problem of exclusion and discrimination of other (bi and trans) groups. This experience of double exclusion from heteronormative society and from lesbian and gay communities was, however, a catalyst for the bisexual movement to emerge.

Taking the lessons and principles learned in lesbian and gay movements, bisexual people began to talk openly about constraints they were facing in the society and in queer communities. Bisexual activists began to press lesbian and gay activists to include bisexuality in the names of pride marches, cultural events, and other activities. By the early 1990s, many networks of formal and loose groups had organized, providing the background for bisexuality to consolidate itself as a core identity for many.

It was also the time when queer studies began to flourish, and it was there, although again not without problems, where bisexual activists and scholars invested a lot of energy, hopes, and interests.

Transsexual and Transgender Movement

To begin with, it may be useful to define the subtle yet important difference between *transsexual* and *transgender*. The first one refers to persons who perceive themselves mentally, psychologically, and emotionally as being of the opposite sex and gender to those of their born-with bodies. Transsexual people will ultimately go through transition and with the help of medical treatment and surgical body shaping, will become of the sex and gender they always felt to be. Conversely, transgender people, although perceiving themselves as of opposite gender, may not necessarily wish for surgical sex-change operations (although it is not an excluded option). Hence, a transgender person may choose to adopt only partial treatment (e.g., hormones) and be passing for the other gender yet still technically living with inborn genitals. The notion of transgender is therefore wider and encompasses transsexual as one of the ways of living trans life. The abbreviation FTM means *female to male*, and MTF means the opposite.

As in the case of lesbian, gay, and bi movements, first traces of transgender activism can be found in the late-19th-century sexology, most notably in writings of Karl Heinrich Ulrichs. He conceived of the homosexual people as “Urnings”—people of male sex with female souls, and vice versa. Thus, clearly, the ideas of homosexuality, bisexuality, and transgenderism were intertwined together. It is, however, the mid-20th century that witnessed the first more organized (in the modern sense) attempts at building transgender community.

In 1952, Virginia Prince, a transgender women (MTF), set up a first magazine: *Transvestia: The Journal of the American Society for Equality in Dress*. It was relaunched

in 1960 and established itself as a long-lasting core of a newly forming transgender peer network. What followed was a slow but consequent establishment of support, friends, and information networks, with more informal groups becoming more visible and active in the public sphere. A good example is the second half of the 1960s in the San Francisco area. A cooperation between city council officials, transgender activists, and health professionals gave rise to a well-established network of support services (like easy access to hormonal therapy, psychological counseling, etc.) and education.

Another symbolic moment in the history of transgender movement is, already mentioned, Stonewall Inn riots in St. Christopher Street, New York, 1969. This date is usually ascribed to the beginnings of the gay and lesbian liberation movements. It is, however, worth remembering that Stonewall bar was actually a commonplace of gathering for transgender people and drag queens (male, usually homosexual, impersonators of femininity). Nonetheless, soon during the decade of 1970s, the transgender community got disillusioned and disappointed with gay and lesbian liberation on the one hand, and the developing second wave of feminism, on the other.

As it was the case with bisexuals, transgender people were also often excluded from both movements. Many feminists of that time felt that FTM men were traitors of the cause, or MTF women were seen as not real or as invaders or perpetrators of the feminine sphere. Lack of understanding of transgender people's problems was also clear in ranks of lesbian and gay liberation groups. These often portrayed transgender people as having false consciousness, repeating and strengthening social gender order—rather than trying to dismantle it.

This history of exclusion and hurt became later an important influence on the formation of the queer activism. Around the 1980s, a stronger presence of FTM communities became noticeable, with many FTM social networks emerging across the United States. A certain shift in sexual politics could be again noticed around the second half of the 1980s. It was an alteration in the HIV and AIDS politics, which started to move away from an identity-based type of politics (dominant at the beginning of the epidemic). During this time, transgender, bisexual, lesbian, and gay communities came together, forming queer activism. The relationship between transgender activism and queer approaches is not unproblematic, though.

Some transgender communities feel that queer theory remains nothing but theory and does not meet pragmatics and problems of material life. For them, the term *transgender* became an empty signifier, a theoretical figure of academic queer theorizing, much divorced from the everyday needs of transgender people.

However, it should also be noted that many other trans people felt that it was queer and its stress on the flexibility of identities was where they finally could find their place, not being pinned down to only one category.

Such conflict, present in every movement, is a good example of problems faced by all identity- and non-identity-based movements. It is not a sign of weakness, but rather of an ongoing process, which cannot be enclosed in one definition or a model.

LGBT Politics: Main Issues

After a historical overview, this section takes a conceptual form, presenting the reader with a selection of the most persistent and important issues for the LGBT movements in the last 50 years of the 20th century. These are education and representation, legal changes, same-sex partnerships, and health. The contested nature of these issues is addressed in the “Queer Approaches” section that follows.

Education and Representation

Homophobia spreads because of fear of the unknown. Therefore, one way of counteracting homophobia is to make homosexuality known, through education.

Education as a component of lesbian and gay struggle was usually two dimensional: directed at the homosexual community (consciousness raising, safer-sex information, support groups, psychological help, etc.) and at the majority of society (cultural events like film screenings, festivals, public campaigns, marches, parades, appearance in television and radio programs, etc.). Education may also be seen as the general framework for lesbian and gay politics, not only as one of its tools. Indeed, politics itself can be defined as a form of education focused not only on transmitting existing structures, agents, and institutions but also on transforming them.

It is from this angle, of education as a framework, that visibility (representation) may be treated as part of a larger project (politics); nevertheless, it is often presented as an autonomous strategy. The usually grounding conviction behind its importance is that homosexual people are wrongly portrayed and seen in society as abnormal, sexually vulgar, oversexualized, promiscuous, gender dysfunctional, and so on. Consequently, the task of the LGBT struggle is to reverse these misconceptions by disseminating positive representations of the LGBT community in all its diversity. Thus, for example, the homophile movement suggested gender-normative behavior to their followers, and the gay rights approach stressed professionalism.

Reclaiming History

Reclaiming history is yet another popular strategy for educating societies and generating more positive representations of homosexual people. By calling on famous homosexual figures from the past (often beginning with ancient Greece and Rome, through Michelangelo and Leonardo da Vinci, to various other people from past and present),

LGBT communities are trying to regain the voice, become agents, and write or tell their own histories. As many studies show, the question of whose voices are heard and how stories are told are the central issues in the processes of the creation of social and political inequalities. So reclaiming history would be an act of reappropriating agency by LGBT groups and fighting back against the heteronormativity of sociopolitical space.

Coming Out

“Coming out”—the act of public disclosure of one’s homosexuality—is often seen as a necessary condition and element of LGBT identity politics. It stresses the need for education and visibility by using one’s own voice. The underlying assumption is that dominant discourses are heteronormative and that all nonheterosexual people are subjugated to and oppressed by its regulatory mechanisms. Consequently, the public declaration of one’s own non-heterosexuality is a way to break out from the matrix of domination, regain one’s own voice, and become an agent of one’s own creation. The act of coming out is of crucial importance for LGBT identity politics. In the private act of coming out into the public domain, the individual sexual identity becomes de facto a political statement. Thus, the collective identity is born: As such, homosexuality enables agency and the historical possibility of the emergence of LGBT identity politics.

Pride Parades and Marches

Gay Prides, as they are popularly called, are without doubt the most popular strategy used by the LGBT community. The idea originated in the memorial marches organized to commemorate the Stonewall riots in 1969. Gay Pride stresses the need for public visibility of LGBT subjects. As the name suggests, the aim was to show society proud homosexuals, individuals that were actively pursuing their liberties, rather than shamefully and secretly seeking social accommodation. It was clearly a liberationist stance, opposing the more tame homophile attitudes of the 1950s and 1960s. Parades take the feminist idea that the personal is political quite literally and bring sexuality onto the streets. Pride marches have evolved through time and changes in geopolitical context. From early political riots (like those around Stonewall and other places), through political manifestations (like marches on Washington), to apolitical, carnivalesque entertainment (prevailing nowadays), what they have in common is their embodiment of the 1970s slogan: “Out of the closets and into the streets.”

Homophobia, Racism, and Androcentrism

Another profound contribution of feminist and LGBT scholars is to put a stress on intersectionality. Intersectional

analysis means to look at social and cultural realities through a prism of no one single factor but many interrelated factors—or to see society and politics as inevitably intertwined. Therefore, it is suggested that to understand homophobia, we must also look at racism and androcentrism in society. Intersectionality means to look at class, gender, sexuality, race, ethnicity, religiosity, and other factors as always already intertwined. Hence, to understand one, we need to grasp the dynamics of the other aspects as well.

This was originally suggested by feminist and LGBT activists of color, who often felt that both movements were not properly dealing with their more specific problems, for example, racism or immigration control. Intersectionality was later adopted in queer approaches as the crucial method of understanding and dealing with sociopolitical reality.

Legal Change

LGBT politics during the decades of the 1950s to 1970s were very much focused on the cultural and social aspects determining discrimination against homosexual people. By the end of the 1970s, with the growth of the gay rights approach, more attention was given to the legal framework. Three major areas of interest may be distinguished: decriminalization of homosexuality, imposing antidiscrimination regulations, and providing legal regulations of same-sex partnerships.

Although not on a national level, some individual states upheld so-called sodomy laws, penalizing (mostly male) same-sex sexual behavior. These laws are seen by LGBT groups as belonging to the previous epoch and violating basic human rights. Therefore, eradication of these regulations became an important part of LGBT struggle in the 20th century.

On the other hand, more positive laws were suggested as worth inscription into the legal code, especially regulations that would directly mention sexual orientation as one of the grounds on which discrimination is prohibited. It is argued that current existing (if at all) antidiscrimination laws are hard to execute if no direct inscription of sexual orientation is present in those bills.

Finally, the legal recognition of same-sex partnerships has become one of the most important issues in LGBT politics. For that reason, it is discussed next as a separate section.

Same-Sex Partnerships

Same-sex partnerships and gay marriages became the core of LGBT politics in America and worldwide in the second half of the 1990s. The distinction between legally recognized partnerships and marriages is not only linguistic but also has symbolic meaning. The use of the term *gay marriage* denotes not only the plea for legal regulation of

homosexual relationships but also highlights inequalities in the social regulation of homo- and heterosexuality. As such, advocates of gay marriage do not want “just like marriage” same-sex partnerships, because this still dialectically places heterosexual marriage as the ideal. But rather, the use of the term *gay marriage* is an attempt at transforming wider social relations of privilege, emphasizing existing customs and habits of power. It is to make social institutions as marriage defined as not exclusively heterosexual.

Why has the struggle for legally recognized homosexual relationships become so important? There are various reasons. One of them is the growing dominance of gay rights, among other approaches, in LGBT politics throughout the 1980s and 1990s. This tactic intensively deploys the legal-structural framework for the advancement of rights for LGBT people. It is driven by a belief that legal adjustment will erase homophobia and discrimination in society (a change from what was previously discussed).

Another reason is connected to the wider developments in North America and other Western liberal democracies. Arguably, the 1990s brought about the amelioration of living conditions of Western societies and, after the 1980s backlash, the liberalization of social attitudes. Some queer-oriented critiques point toward the fact that the aspiration of gay people toward same-sex partnerships reflects neoliberal consumerism and a general adaptation to middle-class standards and norms. On the other hand, it is also seen as the obvious next step on the way to equality. After fighting for negative freedom from discrimination, same-sex partnerships would signify positive freedom to self-realization and fulfillment. Same-sex partnerships remain today one of the most contested and hotly debated areas of LGBT politics.

Health

The issue of medical discourses, the health care system, equal access to treatment, and governmental spending on health-related research is the last theme of LGBT politics. It is possible to distinguish three areas of interest: medical discourses about sexuality, HIV- and AIDS-related issues, and transsexual people's needs.

Pathologizing discourses, which present homo-, bi-, and transsexuality as aberrations and deviancy, were strong in various medical disciplines, most clearly in psychology and psychotherapy. From the 1950s onward, the reversal of those stigmatizing discourses was of primary importance for LGBT movements. Strongly related to the issue of representation, finding another way of talking about sexuality in medicine (a way that would not pathologize its nonheterosexual forms) took various forms in various historical moments.

The decade of the 1980s made clear that the governmental response to the HIV and AIDS pandemic was totally inefficient in facing and combating its spread.

There was a need of new treatments, funding research, experimental approaches, and providing help and support to those already affected (both patients and their relatives). LGBT groups also stressed the need for preventive policies. They lobbied for outreach programs (e.g., distribution of free condoms) and better sexual education in schools.

Also, the rise of the trans movement made the improvement of health services in the United States one of the major concerns. The often long-haul needs of transsexual people during their transitions highlighted such problems as more open access to psychological help, as well as medical and surgical treatment. The emphasis was put on public funding of those services, simplification of bureaucratic procedures, and depathologization of medical discourses around transsexuality and transgenderism.

Queer Approaches

First, it has to be noted that the term *queer* is a highly contested area, with significant differences in use and understanding of terms and attitudes. Since it is such a contested area, the way it is presented here is just one of many articulations. It is usually welcomed as celebrating diversity and fluidity (as opposed to sometimes too-rigid LGBT strategies and definitions), incorporating greater awareness of global political processes. However, queer approaches have also been criticized by feminist, feminist of color, and LGBT scholars and activists for rendering political activism less possible, paying too much attention to discursive practices and not enough to material conditions of living, and for being too theoretical and lacking empirical inquiry. Most recent queer writing indicates that these discussions have been beneficial for the ongoing development of queer studies. For different ways of synthesizing, please see the rich works of, for example, Michael Warner, Arlene Stein, and Ken Plummer.

Critique of Identity Politics

Although LGBT politics took various shapes and forms in the course of its evolution, the issue of identity politics can be said to be at its most prominent between 1960 and 1980. Without entering into a nuanced debate about identity politics as such, LGBT identity politics can be summarized as follows. It is about treating sexuality and gender as the dominant, core essence of a person's identity. Hence, we talk about gay identity or bisexual identity. LGBT identity politics has been influenced by other identity politics movements from the American counterculture of the 1960s and 1970s. This approach presumes that all homosexual (and for that matter bisexual and transsexual) people experience the same sort of oppression and that sexuality is the most important characteristic for their sense of self.

This position is often referred to as essentialism—because, of many aspects that create our identity, only one

is seen as the most important, essential, to understand a person. LGBT identity politics is thus an essentialist strategy that (over)values sexuality, placing it above class, race, gender, cultural affiliations, and other categories. This strategy proved at times to be a successful and efficient way of challenging homophobia and discrimination but also has its clear drawbacks.

Among other things, queer movements and queer academic scholarship point to the constraints imposed by identity politics, proposing alternative solutions in activism and the way people think sex. As already pointed out in the case of bisexual and transsexual movements, LGBT identity politics often lead to the hegemonizing of lesbian and gay people over bisexuals and transsexuals. The very mechanism of exclusion that homosexual communities fought against were often redeployed by the same groups against their own minorities of bisexual and transsexual people.

Queer activism of the late 1980s and 1990s—as represented in such groups as ACT UP, Lavender Menace, or Queer Nation—was also characterized by a return to more in-your-face types of activism. These approaches were similar to those that were widely adopted at the early stages of gay liberation and are in stark contrast to more formally structured lobbying strategies used by pressure groups of the gay rights approach.

Critique of Politics of Visibility

Following critical attitudes toward identity politics, similar critical stances are developed toward politics of visibility.

Queer approaches scrutinize the act of coming out, a favored strategy of LGBT identity politics. Inspired by the work of French philosopher Michel Foucault, who showed how unequal power relations are established in any act of confession, the queer approach refuses coming out as an act that ascertains heterosexuality at the core of its politics. It is suggested that such an act makes LGBTQ people subjects of surveillance within heteronormative (see subsequent definition) social regulations. Directly connected to this issue is the question of representation.

The importance given by LGBT groups to public visibility has been critically reassessed by queer activists. The question is whether the public (under)representation is always already a bad and negative thing for the LGBT community. Queer approaches argue that strong pressure for more visibility has normalizing effects, excluding non-normative (queer) representations. Media inclusion may be potentially exploitative of those members of community, who do not fit into standardized vision of a normal LGBT person.

Critique of Neoliberalism

Queer politics is also preoccupied with the impact of consumerism and neoliberalism on the creation of a gay lifestyle. It is noticed that consumerist practices and

entertainment encouraged by capitalism promote white, male, middle-class lifestyle, at the expenses of other social groups. Alexandra Chasin (2000), Lisa Duggan (2002, 2004), and Lauren Berlant (1997) show how neoliberal capitalism locates its interests unequally and how in neoliberal democracies citizens are treated more as consumers than anything else. This process only strengthens inequalities, reproducing the social core (usually white and male) and the margins (women, LGBT people, immigrants, and people of color). Cultural practices of consumerism, like gay tourism or sex tourism, were shown to work alongside the economic dominance of the global North over global South. The interest in material rudiments of political and social position occupied by LGBTQ people has developed in response to critical voices, which rightfully pinpointed that literary and cultural origins of queer theory obscured the queer analysis of everyday life in its material, not only discursive, dimensions.

The aforementioned queer critiques build up toward the larger queer project of challenging heteronormativity of public and private life. The term *heteronormativity* could be defined as a set of overarching practices, discourses, and arrangements facilitating domination of reproductive heterosexuality as the normative category of modern societies, organizing people according to their genders, social positions, ethnic markers, and sexual choices.

Geopolitics: Neoimperialism, Terrorism, and Nationalism

Another important development within queer studies is the work on the geopolitical and temporal practices of U.S. governments and Western and U.S. LGBT activism and scholarship. Book editors such as Martin Manalansan, Cindy Patton, Benigno Sánchez-Eppler, Arnoldo Cruz-Malavé, and John Hawley have contributed to the quickly growing field of inquiry combining queer and postcolonial theories. Authors gathered in their books offer crucial insights into such processes as citizenship status of queer migrants and queer people of color, sexuality and ethnicity in diasporas, asylum policies and sexuality, sexuality and nationhood, control of bodies, and national borders.

Globalization of LGBT identities and the global gay movement are also important issues of study. Especially the spread of the Western and U.S. models of LGBT activism around the world as universal is questioned. As in the case of gay tourism, Western LGBT activism is analyzed in the light of globalization and cultural hegemonic practices of the West and the United States. The standardization of sexualities and their use for wider political prospects is scrutinized, especially in relation to American nationalism, terrorism, and neoimperialism.

For example, Judith Butler (2009), a key figure of queer studies, and Jasbir Puar (2007) show how some acceptance of gay rights has been adopted in American nationalist discourse as a sign of progress. In turn, this has been used to legitimize the war on terror's "civilizing mission" in the

Middle East. However controversial, queer approaches' contribution was to show how racialisation of the otherness (in the international relations) is intrinsic to certain forms of normalization of homosexuality in domestic politics.

Conclusion

The impact of the LGBT communities on the political sphere continues to be diverse and come from various angles, from social movements embracing identity politics and focusing on cultural politics of societal change to governmental pressure groups lobbying for legal change. It ranges from street activism, through academic theorizing, reaching to influence modern understandings of political ideologies.

Of the various issues raised throughout history, those of education, health, racism, exclusion, androcentrism, ideologies, nationalism, hope, AIDS, ethics, intimacies, and visibility, among others, were dealt with.

In a sense, the success of LGBT and Q politics is also seen through the existence of this very chapter. Political science was rather reluctant to incorporate sexuality as a prominent aspect worth theorizing. Dominated by the positivist epistemology (philosophical perspectives on what constitutes knowledge) of so-called objective truth, in search of universal models and theories explaining voting behavior, diminishing of party politics, for example—there was no space for sexuality. However, the liberal ideology, which has greatly contributed to this exclusion, has changed, as did other ideologies. Classical liberal thought put sexuality into the private sphere, which, according to John Locke, should not overlap with the political domain. It is very much thanks to the feminist and LGBT struggles, which insist that private is political, that nowadays the understanding of politics is wider and more inclusive. And that political science handbooks more often do include chapters like this one.

Finally, the chapter could be summarized under the heading "sexual citizenship." How governments are dealing with sexuality, gender, citizenship, and national and ethnic identities are pressing issues that secure more and more attention. Hence, the term may prove useful when trying to understand most recent history. Embracing tolerance of homosexuality as a national feature, using it to control immigration, or to mobilize international support for imposing legislatures to prevent discrimination are just a few examples of how LGBT issues have been used so far in 21st-century politics.

References and Further Readings

- Adam, B. (1997). *The rise of a gay and lesbian movement* (Rev. ed.). New York: Twayne.
- Adam, B., Duyvendak, J. W., & Krouwel, A. (1999). *The global emergence of gay and lesbian politics*. Philadelphia: Temple University Press.

- Bell, D., & Binnie, J. (2000). *The sexual citizen: Queer politics and beyond*. Cambridge, UK: Polity Press.
- Berlant, L. (1997). *The queen of America goes to Washington City: Essays on sex and citizenship*. Durham, NC: Duke University Press.
- Binnie, J. (2004). *The globalization of sexuality*. London: Sage.
- Blasius, M., & Phelan, S. (Eds.). (1997). *We are everywhere: A historical sourcebook of gay and lesbian politics*. New York: Routledge.
- Brown, W. (2006). *Regulating aversion: Tolerance in the age of identity and empire*. Princeton, NJ: Princeton University Press.
- Butler, J. (2009). *Frames of war: When is life grievable?* London: Verso.
- Chasin, A. (2000). *Selling out: The gay and lesbian movement goes to market*. New York: St. Martin's.
- Cooper, D. (1995). *Power in struggle: Feminism, sexuality and the state*. Buckingham, UK: Open University Press.
- D'Emilio, J., & Turner, W. B., & Vaid, U. (Eds.). (2000). *Creating change: Sexuality, public policy, and civil rights*. New York: St. Martin's.
- Duggan, L. (2002). The new homonormativity: The sexual politics of neoliberalism. In R. Castronovo & D. Nelson (Eds.), *Materializing democracy: Toward a revitalized cultural politics* (pp. 175–194). Durham, NC: Duke University Press.
- Duggan, L. (2004). *The twilight of equality: Neoliberalism, cultural politics, and the attack on democracy*. Boston: Beacon Press.
- Eaklor, V. L. (2008). *Queer America: A GLBT history of the twentieth century*. Westport, CT: Greenwood Press.
- Hennessy, R. (2000). *Profit and pleasure: Sexual identities in late capitalism*. New York: Routledge.
- Hertzog, M. (1996). *The lavender vote: Lesbians, gay men, and bisexuals in American electoral politics*. New York: New York University Press.
- Kaplan, M. (1997). *Sexual justice: Democratic citizenship and the politics of desire*. New York: Routledge.
- Kulpa, R., & Mizelińska, J. (Eds.). (2010). *De centring Western sexualities: Central and East European perspectives*. London: Ashgate.
- Manalansan, M., & Cruz Malavé, A. (Eds.). (2002). *Queer globalizations: Citizenship and the afterlife of colonialism*. New York: New York University Press.
- Meyerowitz, J. (2002). *How sex changed: A history of transsexuality in the United States*. Cambridge, MA: Harvard University Press.
- Parker, R., Barbosa, R. M., & Aggleton, P. (Eds.). (2000). *Framing the sexual subject: The politics of gender, sexuality and power*. Berkeley: University of California Press.
- Patton, C., & Sánchez Eppler, B. (Eds.). (2000). *Queer diasporas*. Durham, NC: Duke University Press.
- Puar, J. (2007). *Terrorist assemblages: Homonationalism in queer times*. Durham, NC: Duke University Press.
- Rich, A. (1993). Compulsory heterosexuality and lesbian existence. In H. Abelove, M. A. Barale, & D. M. Halperin (Eds.), *The lesbian and gay studies reader* (227–254). New York: Routledge.
- Richardson, D. (2000a). Constructing sexual citizenship: Theorizing sexual rights. *Critical Social Policy*, 20(1), 105–135.
- Richardson, D. (2000b). *Rethinking sexuality*. London: Sage.
- Rimmerman, C. (2002). *From identity to politics: The lesbian and gay movements in the United States*. Philadelphia: Temple University Press.
- Stevens, J. (1999). *Reproducing the state*. Princeton, NJ: Princeton University Press.
- Tucker, N. (Ed.). (1995). *Bisexual politics: Theories, queries, and visions*. Binghamton, NY: Harrington Park Press.
- Williams, W. L., & Retter, Y. (Eds.). (2003). *Gay and lesbian rights in the United States: A documentary history*. Westport, CT: Greenwood Press.
- Winnubst, S. (2006). *Queering freedom*. Bloomington: Indiana University Press.

INDEX

Main topics and page numbers are in **bold**.

- A priori (abstract) judicial review, **1:188**
- A priori posteriori (concrete) judicial review, **1:188**
- Absolute monarchy, **1:202**
- Accommodationism, **2:833–834**
 - religion and, **2:852–853**
- Accountability
 - bureaucracy and, **2:748**
 - democratic, **1:174–175**, 268
 - of nonstate actors, **1:419**
- Acculturation, **2:833**
- Acid rain, **1:445**
- Actual idealism, **2:642**
- Actual surplus, **1:102**
- Adams, John, **2:690–691**
- Adams, Sam, **2:690–691**
- Additional member system, **1:160–161**
- Administrative man, **1:35–36**
- Adorno, Theodor, **1:55**, 252
- Adverse selection, **1:43**, 44
- Advertising, political, **2:710**
- Advocacy coalition framework, **2:793**
- Afghanistan
 - democracy and, **1:214**
 - nationalism in, **2:637**
 - post-Taliban elections, **1:164–165**
 - war on terrorism and, **2:824**
- Africa
 - communism in, **2:674**
 - economic policy and, **1:47**
 - See also individual country*
- African Americans, **2:836–837**
- Agonistic pluralist theory, **2:611–612**
- Agrarian parties, **1:151**
- AIDS, **1:237**, **2:858**, 861–862
- Al Qaeda, **1:118–119**, 120, 121, 122, **2:824**
- Albania, **1:145**
- Albany Plan of Union, **2:687**
- Al-Farabi, **2:570**, **2:572**
- Algeria
 - Mediterranean Action Plan, **1:427**
 - rentier effects in, **1:398–399**
- Alighieri, Dante, **2:582–583**
- All courts model of judicial review, **1:188**
- Allende, Salvador, **1:178**, 182
- Allison, G. T., **1:36–37**
- Allison, Graham, **2:790–791**, 826
- Al-Mahdi, Muhammad, **2:571**
- Almond, Gabriel, **1:6**, 39, 71, 72, 75, 76, 83, 93, 201–202, **2:753–754**, 757
- American Political Science Association (APSA), **2:605**
- Anabaptists, **2:585**
- Analysis of variance (ANOVA), **2:499**
- Analytic structures, **1:78**
- Anarchism, 2:625–632**
 - anarchist feminism, **2:629**
 - classic thinkers on, **2:626–627**
 - common ideas in, **2:630–631**
 - contemporary variants, **2:627–629**
 - feminist, **2:629**
 - ontological anarchism, **2:628–629**
 - philosophical anarchism, **2:630**
 - postanarchism, **2:629**
 - primitivism, **2:628**
 - security dilemma and, **1:378**
 - social ecology, **2:627–628**
 - terrorism, **1:117**
 - theory and practice, **2:630–631**
- Ancients, 2:553–559**
 - authenticity problem, **2:557**
 - common ground among, **2:554–556**

- human nature, 2:555–556
- interpretation problem, 2:557
- nature, 2:554–555
- next step, 2:558–559
- polis, 2:556
- study of ancient political philosophy, 2:556–559
- truth problem, 2:557–558
- Anderson, Benedict, 2:635
- Androcentrism, 2:860–861
- Angell, Norman, 1:321
- Anomic interest groups, 2:753
- Anticolonial nationalism, 2:636–637
- Antiessentialism, 2:629
- Anti-federalism, 2:695
- Antiglobalization, 2:627–628
- Anti-Malthusian, 1:442–443
- Antirealism, Duverger's law and, 2:453
- Anti-Socialist Laws, 2:656
- Appellate courts, 1:189–190
- Aquinas, Thomas, 2:581–582, 588
- Arab nationalism, 2:636–637
- Arab socialism, 2:677
- Arendt, Hannah, 1:252
- Argentina, decentralized education in, 1:174
- Aristotle, 1:61, 62, 75, 94, 269, 2:554, 555, 556, 558, 559, 588
- Arrow, Kenneth, 1:37, 273, 2:537, 608
- Arrow's paradox, 2:608, 2:609
- Art sabotage, 2:628
- Articles of Confederation, 2:689–690
- Aryan myth, 2:642–643
- Asia
 - adoption of communism in, 2:674
 - democratic breakdown in, 1:279
 - See also* Asian political thought; *individual country*
- Asian political thought, 2:560–567
 - classical Asia, 2:561–562
 - colonial Asia, 2:562–563
 - cultural grounding of concepts, 2:565–566
 - modern Asia, 2:564–565
 - theoretical approach, 2:560–561
- Asian values, 1:254
- As-if-random conditions, 2:527–528
- Assimilation, 2:833
- Associational interest groups, 2:753
- Atheism, 2:850
- Athens, democracy in, 1:268–269
- At-large elections, 2:704
- Atlas Shrugged* (Rand), 2:618
- Atomization, 1:20
- Attitudes, political, gender effect on, 2:841
- Attitudinal model
 - of judicial behavior, 2:735–736
 - of judicial decision making, 1:187–188
- Attribution theory, 1:54
- Augustine (Saint), 2:579–580
- Australia
 - as federal system, 1:169
 - high courts in, 1:188, 190
 - two-and-a-half party system in, 1:154
- Austria, as federal system, 1:169
- Austrian School, 2:618
- Authoritarian personality, 1:55
- Authoritarian states, types of, 1:94–95
 - bureaucratic-authoritarian system, 1:94
 - corporatist state, 1:94
 - neopatrimonial state, 1:94–95
 - patron–client relationship, 1:95–96
 - posttotalitarian state, 1:95
 - totalitarian state, 1:95
- Authoritarianism. *See* **Semiauthoritarianism; Totalitarianism and authoritarianism**
- Autonomy
 - bureaucracy and, 2:749
 - of the state, 1:92
- Averroes, 2:573
- Avicenna, 2:573
- Ba'athism, 2:637, 677
- Bacon, Francis, 2:589, 687
- Bagdikian, B., 2:709
- Baker v. Carr*, 2:782
- Bakunin, Mikhail, 2:626, 649
- Balance of power, 1:361–367**
 - bandwagoning and, 1:363
 - collective security and, 1:363
 - core concepts/dynamics of, 1:361–363
 - emergence in practice, 1:363
 - French threat in Europe, 1:364
 - in nuclear age, 1:365–366
 - in 20th century, 1:364–365
 - institutionalization of, 1:363–364
 - origins of, 1:363–366
 - polarity and, 1:362–363
 - soft balancing, 1:366
 - unipolar system, 1:366
- Ballot structure, 1:160
- Bandwagoning, 1:363
- Baran, Paul, 1:101, 102, 103
- Barber, James, 1:355–356
- Barres, Maurice, 2:640
- Baudrillard, Jean, 1:16, 19

- Bay of Pigs, **2:677**
See also Cuban missile crisis
- Beard, Charles A., **2:695–696**
- Bebel, August, **2:655, 2:657**
- Beblawi, H., **1:335–336**
- Begin, Menachem, **1:354**
- Behavioral economics, **1:52**
- Behavioral revolution, as root of neoinstitutionalism, **1:22–23**
- Behavioralism, **1:3–4, 1:6–8**
 “Consensus and Ideology in American Politics,” **1:6–8**
 definition and overview, **1:6**
 major tenets of, **2:464–465**
 political science and, **2:453–454**
 Strauss’s criticism of, **1:61**
- Belarus, **1:340–341, 355**
- Belgium, **1:146, 174**
- Belief sampling, **2:516**
- Bell, D. A., **2:565–566**
- Bentham, Jeremy, **1:129, 320, 2:595, 828**
- Berelson, B. R., **2:814**
- Berger, Peter L., **1:211**
- Berns, W. **1:66**
- Bernstein, Eduard, **2:656, 657–658, 661**
- Bertalanffy, Ludwig von, **1:71**
- Besci, Z., **1:134–135**
- Bey, Hakim, **2:628**
- Bias in media, **2:709–710**
- Bicameralism, **1:169**
- Big shuffle concept, **1:96**
- Bill of Rights, **2:696–697**
- bin Laden, Osama, **1:118–119, 120, 122, 2:824**
- Biocapacity, **1:135**
- Biological determinism, **2:642–643**
- Bipartisan Campaign Reform Act, **2:776–777, 802**
- Bisexual movement, **2:858–859**
- Bivariate regression, **2:478–481**
- Black, Bob, **2:627–628**
- Blackmail potential party system, **1:154**
- Blair, Tony, **1:339**
- Bloom, Allen, **1:63, 66**
- Bloomington school, **1:39**
- Blumenthal, Sidney, **2:797**
- Bodin, Jean, **2:585, 2:589**
- Bolsheviks, **2:667–668**
- Bonding social capital, **1:195**
- Bookchin, Murray, **2:627, 2:628–629**
- Bosnia, **1:215, 2:637**
- Bounded rationality, **1:35–37, 52, 380**
- Bourgeois democracy, **2:652**
- Bourgeoisie, **2:652, 657, 665**
 tolerance of, **2:678**
- Brazil, **1:161, 169, 174**
- Breakthrough coup, **1:126**
- Breuning, M., **2:492**
- Brezhnev, Leonid, **2:669**
- Bridging social capital, **1:195**
- “Bringing state back in” concept, **1:91, 92–93**
- Brown, L. Susan, **2:629**
- Brown v. Board of Education*, **2:701, 833, 836**
- Bryce, James, **1:3**
- Buchanan, James, **1:37–38, 40**
- Buddhism, **2:566**
- Budge, I., **1:156**
- Bureaucracy, American**
 accountability issues, **2:748**
 American foreign policy and, **2:826–827**
 applications and empirical evidence, **2:746–748**
 autonomy issues, **2:749**
 critique of, **2:746, 747**
 defining bureaucracy, **2:743–744**
 development of, **2:744**
 early scholarship on, **2:744–745**
 empirical studies on, **2:748**
 future research issues, **2:749–750**
 organization of modern bureaucracy, **2:745–746**
 policy and, **2:747, 748–749**
 radical comprehensive approach to, **2:747**
 scientific management theory on, **2:747**
 testing theories of, **2:747–748**
 U.S. Congress and, **1:48, 2:721–722**
- Bureaucratic efficiency vs. individualized justice, **1:13, 19**
- Bureaucratic politics model (BPM), **1:340**
- Bureaucratic-authoritarian system, **1:94, 253**
- Burke, Edmund, **2:593–594, 621**
- Bush, George H. W., **2:728, 774**
- Bush, George W., **1:179, 182, 340, 353–354, 2:761, 775, 824**
- Cabinet staffing, by president, **2:727–728**
- Calhoun, John, **2:760–761**
- Calvin, John, **2:584–585, 2:587–588**
- Campaign advertising, **2:800–801**
- Campaign effects, **2:802**
- Campaign financing, **2:776–777**
- Campaigns, 2:796–804**
 advertising, **2:800–801**
 campaign effects, **2:802**
 fund-raising, **2:797, 801–802**
 gender differences in, **2:800, 843**
 levels of, **2:796–800**
 negative campaign ads, **2:710, 801**
 partisan, **2:799–800**

- policy implications, 2:802
- political consultants, 2:798–799
- professionalism in, 2:799
- state-level, 2:798
- theory on, 2:796–797
- top-of-the-ticket contests, 2:798
- Campbell, A., 2:817–818
- Canada
 - decentralized education in, 1:174
 - ethnic conflict in, 1:143
 - as federal system, 1:169, 172
 - financial crisis in, 1:402
 - high courts in, 1:188, 190
 - NAFTA and, 1:426–427
 - party capability thesis and, 1:190
 - regionally concentrated ethnic groups in, 1:146
 - two-and-a-half party system in, 1:154
- Capital: A Critique of Political Economy* (Marx), 2:648, 649–650, 665
- Capitalism
 - center–periphery structure in, 1:102
 - Marxism on rise and collapse of, 2:652
 - semi-periphery structure in, 1:103
- Cardoso, Fernando, 1:104
- Carey, John, 1:178
- Cartel parties, 1:154
- Carter, Jimmy, 1:354, 2:728, 774
- Carter, Ralph, 2:825, 2:826
- Case studies, 1:293–299**
 - comparative, across-case study, 1:294, 295
 - comparative politics example, 1:298
 - configurative-idiographic study, 1:295
 - crucial case study, 1:295
 - debate within discipline, 1:294–297
 - disciplined-configurative study, 1:295
 - examples of approach, 1:297–299
 - heuristic case study, 1:295
 - individual, within-case study, 1:294
 - international relations example, 1:298–299
 - least-likely case study, 1:295
 - micronumerosity in, 1:296
 - most-likely case study, 1:295
 - structured, focused comparison in, 1:297
 - in urban politics, 1:31
 - U.S. politics example, 1:297–298
- Castro, Fidel, 2:677
 - See also* Cuba
- Catchall parties, 1:153–154
- Categorical ballot, 1:160
- Catholicism. *See* **Christian political thought**
- Caucus for a New Political Science, 1:8
- Census outbidding, 1:220
- Center–periphery, 1:102
- Centralization, 1:169
- Centralpetalism, 1:171
- Centrifugal party system, 1:155
- Centripetal party system, 1:155
- Chaos theory, 1:402
- Chavez, Hugo, 1:263, 264
- Checks and balances, 2:693
- Chicago School, 1:6
- China
 - adoption of communism in, 2:674
 - ethnic conflict in, 1:143
 - Maoism, 2:669–670, 677–678
 - nuclear weapons in, 1:378
 - rivalry formation policy in, 1:382
 - social movements in, 1:231
 - See also* **Asian political thought**
- Chirac, Jacques, 1:182
- Chomsky, Noam, 2:627
- Christian democratic parties, 1:151
- Christian political thought, 2:577–586**
 - Anabaptists, 2:585
 - Aquinas on, 2:581–582
 - Augustine on, 2:579–580
 - Bodin on, 2:585
 - Calvin on, 2:584–585
 - civil disobedience, 2:586
 - contemporary manifestations, 2:585–586
 - Dante on, 2:582–583
 - emergence in classical world, 2:578–579
 - Hooker on, 2:585
 - John of Salisbury, 2:580–581
 - liberation theology, 2:586
 - Luther on, 2:583–584
 - Marsilio de Padua on, 2:583
 - Reformation and, 2:578, 583–584
- Citizenship, defining, 2:659
- City of God* (Augustine), 2:579–580
- City of Man* (Augustine), 2:580
- Civic culture, political culture and, 1:201–202
- Civil law system, 1:187
- Civil nationalism, 2:636
- Civil Rights Act, 2:837
- Civil rights movement, 1:227
- Civil society, 1:193–200**
 - associational life, importance of, 1:194, 198
 - defining, 1:196, 1:268
 - democracy and, 1:196–197
 - lines of research, 1:197–198
 - measurement issues, 1:198–199
 - social capital, 1:193, 194–195
 - social networks, 1:193, 195–196

- technology and, 1:199
- tertiary associations in, 1:199
- traditional vs. alternative, 1:198
- Civil wars, 1:107–114**
 - conceptual definition of, 1:110
 - death totals as metrics for, 1:110–111
 - ethnic conflict, 1:111
 - failed states and, 1:109, 110
 - greed as cause of, 1:112
 - grievances as cause of, 1:112–113
 - internal political violence vs., 1:108–110
 - interstate wars vs., 1:107–108
 - levels of analysis, 1:111–112
 - policy implications, 1:113
 - relative depravity and, 1:111
 - secessionist conflicts, 1:109
 - social revolutions, 1:109–110
 - state-based explanations of, 1:113
 - weak states and, 1:109, 113
- Clash of civilizations, 1:147–148, 206–207, 212
- Clash of Civilizations* (Huntington), 1:212
- Class-action lawsuit, 2:756
- Classic subtype of democracy, 1:260
- Classical constructivism, 2:472
- Classical realism, 1:312, 313–315
- Classical republicanism, 2:687
- Climate change, 1:335–337
- Clinton, Bill, 1:178, 2:706, 728, 774, 824
- Closed list party system, 1:160
- Close-ended questions, 2:515–516
- Coalition potential party system, 1:154
- Coase, Ronald, 1:46
- Coefficient of determination (R^2), 2:481–483
- Cognitive consistence theory, 1:55, 56
- Cognitive dissonance, 1:55–56
- Cohabitation, 1:178
- Cold war
 - democratic peace and, 1:390
 - deterrence theory during, 1:369–370
 - rivalry formation policy during, 1:382
 - Soviet Union and, 1:381–382, 2:637
- Coleman, J. S., 1:76, 194–195
- Collective efficacy, 1:20
- Collective security, balance of power and, 1:363
- Collectivist-civic nationalism, 1:220
- Collectivist-ethnic nationalism, 1:220
- Colonialism, 1:280
 - dependency theory on, 1:329
- Commercial pacifism, 1:320
- Commission form of city government, 2:703–704
- Common law legal system, 1:187
- Common Sense* (Paine), 2:593, 619
- Common sense problem, Straussians and, 1:61–62
- Communal contenders, 1:143
- Communism
 - Marx on, 2:653
 - See also* **Leninism, communism, Stalinism, and Maoism; Marxism**
- Communist Manifesto, The* (Marx & Engels), 2:648, 649, 653, 655, 665
- Communist parties, 1:151
- Comparative, across-case study, 1:294, 295
- Comparative case study, 1:31
- Comparative environmental politics and conflict, 1:242–248**
 - conflict types, 1:244
 - critique of, 1:244–245
 - environmental scarcity theory, 1:243
 - group identity conflicts, 1:245–246
 - insurgencies, 1:244
 - policy implications, 1:246
 - simple scarcity conflicts, 1:244, 245
 - social ingenuity and, 1:246
 - social pressures leading to conflict, 1:243–244
 - technical ingenuity and, 1:246
- Comparative federalism, confederalism, unitary system, 1:168–176**
 - bicameralism, 1:169
 - centralization in, 1:169
 - centralpetalism of, 1:171
 - confederal system, 1:169, 170–171, 172–173
 - decentralization in, 1:169, 171–172, 174
 - democratic accountability in, 1:174–175
 - education issues in, 1:174
 - federal system, 1:168–170, 173–174
 - future research issues, 1:174–175
 - language issues in, 1:174
 - policy implications, 1:173–174
 - territorial system, 1:169–170
 - theory on, 1:169–171
 - unitary system, 1:168, 171–172, 174
- Comparative foreign policy (CFP), 1:338
- Comparative judicial politics, 1:186–192**
 - attitudinal model of judicial decision making, 1:187–188
 - Constitutional/appellate courts, 1:188–190
 - courts and legal system, 1:186–187
 - inquisitorial model, 1:190
 - judicial decision making, 1:187–188
 - judicial independence in comparative perspective, 1:190
 - judicial review in comparative perspective, 1:189–190
 - legal model of judicial decision making, 1:187
 - party capability research, 1:190

- strategic model of judicial decision making, 1:188
- trial courts and juries, 1:190–191
- Comparative methods, 1:283–292**
 - applications, 1:289–290
 - congruence method, 1:288
 - deinductive approach, 1:285
 - equifinality problem, 1:288
 - formal modeling, 1:285, 286–287, 290
 - future research issues, 1:291
 - game theory, 1:287, 288–289, 290
 - goals of research, 1:285
 - inductive approach, 1:284, 290
 - interdeterminacy and, 1:286
 - lagged time-series data, 1:289–290
 - measurement issues, 1:287
 - network analysis, 1:285
 - probit analysis, 1:289
 - qualitative, 1:285, 286, 287, 288, 289
 - quantitative, 1:285, 286, 287–288, 289–290
 - rational choice–driven game theory, 1:285, 286
 - scope of research and, 1:285–286
 - statistical models, 1:286
 - theory and fact mediation in, 1:284
 - theory assessment, 1:287–289
 - theory generation and, 1:284–285
- Comparative political parties: Systems and organizations, 1:150–158**
 - agrarian parties, 1:151
 - blackmail potential party system, 1:154
 - cartel parties, 1:154
 - catchall parties, 1:153–154
 - centrifugal party system, 1:155
 - centripetal party system, 1:155
 - Christian democratic parties, 1:151
 - coalition potential party system, 1:154
 - communist parties, 1:151
 - conservative parties, 1:151
 - decline of parties, 1:156–157
 - ecological parties, 1:151
 - election rules and, 1:154
 - elite parties, 1:153
 - first-past-the-post party system, 1:154
 - fragmentation of party, 1:155
 - functional cleavages and, 1:151
 - future of parties/party system, 1:157–158
 - government formation, 1:155–156
 - ideological distance and, 1:154–155
 - liberal parties, 1:151
 - mass parties, 1:153
 - minimal winning coalition, 1:156
 - moderate multipartism system, 1:155
 - multiparty system, 1:154
 - party families, 1:151
 - party in central office, 1:152
 - party in public office, 1:152–153
 - party on the ground, 1:152
 - party organizations, 1:152–154
 - party system, 1:154–155
 - polarized multipartism system, 1:155
 - political party formation, 1:150–152
 - predominant party system, 1:155
 - regional/ethnic parties, 1:151
 - segmented multipartism system, 1:155
 - socialist parties, 1:151–152
 - territorial cleavages and, 1:151
 - two-and-a-half party system, 1:154
- See also* **Electoral system in comparative perspective**
- Comparative politics
 - example of case study on, 1:298
 - game theory in, 2:547–548
 - neoinstitutionalism and, 1:24
- Comparative statics, 2:536
- Competitive authoritarianism, 1:261, 262
- Complex interdependence and globalization, 1:401–406**
 - complex interdependence, 1:322, 324
 - deglobalization and, 1:402
 - economic globalization and, 1:402
 - environment and, 1:403
 - future conceptual directions, 1:404–405
 - global governance and, 1:402–403
 - global networks, 1:402–403
 - globalization theory, 1:401–402
 - labor insecurity and, 1:403
 - least-developed countries and, 1:401
 - multinational corporations and, 1:401
 - network effects and, 1:402
 - policy implications of, 1:403–404
 - social globalization and, 1:402
 - Westernization and, 1:404–405
- Computer waste, 1:440
- Computer-assisted content analysis, 2:495
- Comte, Auguste, 2:461–462
- Concrete structure, 1:78
- Condorcet (Marquis de), 2:536–537
- Confederal system, 1:169, 170–171, 172–173
- Configurative-idiographic case study, 1:295
- Congress. *See* **U.S. Congress**
- Congruence method, 1:288
- Consensual democracy, 1:271
- “Consensus and Ideology in American Politics” (McClosky), 1:6–8
- Consent, 2:591

- Conservatism, 2:620–622
 See also **Liberalism, conservatism, and libertarianism, modern**
- Conservative nationalism, 2:636
- Conservative parties, 1:151
- Consociational democracy, 1:271
- Consolidated democracy, 1:389
- Constant, Benjamin, 2:595
- Constitution of the United States, 2:690–697
- Constitutional convention, 2:690–691, 762–763
- Constitutional courts model of judicial review, 1:188
- Constitutional development, influences on, 2:685–687
- Constructive vote of confidence, 1:180
- Constructivism, 2:470–477**
 criticisms of, 2:474
 defining, 2:471–472
 empirical research program, 2:473–474
 on ethnic identity, 1:142, 218
 future research issues, 2:474–475
 on international organizations/regimes, 1:427
 on international relations, 1:307
 on nonstate actors, 1:416
 postmodernism and, 1:12
 types of, 2:472–473
- Contagion effects, of ethnic conflict, 1:146
- Containment policy, 2:824
- Content analysis, 2:490–496**
 coding, 2:491
 computer-assisted, 2:493–94, 2:495
 defining, 2:490–492
 design issues, 2:493–494
 forms of, 2:492–493
 future research issues, 2:495–496
 human coders, 2:494–495
 intercoder reliability, 2:494–495
 latent/manifest content, 2:492
 qualitative, 2:492
 quantitative, 2:492
 thematic coding in, 2:492–494
- Contextualizing, 2:558
- Continuous revolution, 2:670
- Contract With America, 2:774
- Convention on Long-Range Transboundary
 Air Pollution (LRTAP), 1:445
- Conventional constructivism, 2:472
- Conventional things, 2:554
- Converse, Philip, 2:828, 829
- Cooper, Richard, 1:33
- Copernicus, 2:588–589
- Corporatist state, 1:94
- Corruption, 1:47
- Cost of reproduction, 2:650
- Cost–benefit analysis, 2:792
- Council-manager form of city government, 2:703
- County government, 2:783–784
- Coups d'état. *See* **Political and military coups**
- Courts
 presidency and, 2:731
 U.S. Congress and, 2:721
 women in, 2:844, 845
 See also **International law; Judicial politics, American**
- Covering law, 1:354
- Criminology
 critical, 1:13
 misunderstandings about crime, 1:13, 20
 new penology, 1:18
 punishment, 1:16
 sex offenders, 1:15
 social construction of crime problems, 1:17
- Critical citizen, 1:207
- Critical constructivism, 2:472–473
- Critical criminology, 1:13
- Critical juncture concept, 1:289
- Critique of Pure Reason* (Kant), 2:594
- Croatia, 1:215, 2:637
- Croly, Herbert, 2:619
- Cross-cultural methods, 2:520–521
- Cross-sectional analysis, 2:498–499
- Crozier, Michel, 2:746
- Crucial case study, 1:295
- Cuba
 adoption of communism in, 2:674
 socialism in, 2:676–677
- Cuban missile crisis, 1:36–37, 354, 2:790
- Cuius regio, eius religio*, 2:634
- Cult of a strong authority, 2:641
- Cult of personality, 2:668–669
- Cultural authoritarianism, 1:255
- Cultural factorism in ethnic conflict, 1:145
- Cultural heterogeneity, 2:635
- Cultural homogeneity, 1:278–279
- Cultural modernization, 1:162–163
- Customary law, 1:434
- Cyber crime, 1:309
- Cyclical majorities preference pattern, 2:608–609
- Dahl, Robert, 1:6, 267, 270, 2:606–607, 699–700, 754
- Damasio, Antonio, 1:57
- Daoism, 2:566
- Darwin, Charles, 2:665–666
- Darwinism, 1:75
- De Man, Hendrik, 2:658
- De Monarchia* (Dante), 2:582–583

De Zwart, Frank, 1:13–14

Decentralization, 1:169, 171–172, 174

Decision making. *See* **Leadership and decision making**

Decision trees, 1:26

Declaration of improbability, 1:189–190

Declaration of Independence, 2:688–689

Decolonization, 1:253

Deconstruction, 1:15

Deep ecology, 1:443–144

Defective democracy, 1:260

Defensor Pacis (Marsilio de Padua), 2:583

Deforestation, 1:440

Deinductive approach, 1:285

Delegative democracy, 1:271–272

Deliberative democracy, 1:272, 2:608

Deliberative polling technique, 2:612–613

Demand scarcity, 1:243

Demobilization hypothesis, 2:710

Democracy

bourgeois, 2:652

civil society and, 1:196–197

ideal of, 2:755

media role in, 2:708–709

role of middle class in rise of, 1:82–83

See also **Democratic peace; Democratization, processes of; Models of democracy**

Democratic accountability, 1:174–175

Democratic autonomy, 1:272–273

Democratic centralism, 2:670

Democratic goods, 1:163

Democratic pacifism (monadic), 1:386

Democratic peace, 1:384–392

cold war and, 1:390

counterarguments, 1:388–390

defining *democracy* and *peace*, 1:388–389

democratic pacifism (monadic), 1:386

democratic peace (dyadic), 1:386–387

developing vs. consolidated democracy, 1:389

initiates of war and, 1:389–390

policy implications of, 1:390–391

realist counterargument to, 1:390

states/sovereignty and, 1:384–386

variations within states, 1:385–386

Democratic peace theory (DPT), 1:320, 325

Democratic revisionism, 2:656–658

Democratic transitions, 2:612

Democratization

third wave of, 1:255

See also **Democratic peace; Democratization, processes of; Models of democracy**

Democratization, processes of, 1:275–282

agency and advocacy, 1:279–280

economic preconditions, 1:277–278

external actors, 1:280

future research issues, 1:280–281

modernization, 1:276–277

policy implications, 1:280

social preconditions, 1:278–279

three waves of democratization, 1:276

timing, sequencing, and politics, 1:279

Democrats, 2:772–275

Demonstration effect, 1:402

Deng Xiao Ping, 2:565

Dependency and development, 1:99–106

dependency theory overview, 1:101

dependent development, 1:104–105

“development of underdevelopment,” 1:101, 102

Lenin and, 1:100–101

Marxist influence on, 1:99–100, 101, 102

modern world-systems and, 1:102–103

structuralism and, 1:103–104

surplus value theory and, 1:102

Dependency and world-systems, 1:327–335

critical evaluation of, 1:332–333

dependency overview, 1:328–331

dependent development, 1:330–331

empirical evidence for, 1:333

international trade, 1:410

neocolonialism and, 1:329

policy implications, 1:333–334

theory, 1:328–332

world-systems overview, 1:331–332

Dependent development, 1:104–105, 330–331

Deracialization, 2:838

Derrida, Jacques, 1:14–15

Descartes, René, 2:589

Descriptive representation, 2:841–842

Determinism, economic, 2:650–652, 665

Deterrence theory, 1:368–375

controversies, 1:371–374

defining deterrence, 1:368–369

development during cold war, 1:369–370

future research issues, 1:374

game theory and, 1:372–373

methodology issues, 1:371–374

post-cold war scholarship, 1:370

post-September 11 scholarship, 1:370

rational deterrence, 1:369–370, 371–372

reputation theory and, 1:373–374

Deutsch, Karl W., 1:71, 1:82

Developing democracy, 1:389

Developing world. *See* **Socialism in developing world**

- “Development of underdevelopment,” 1:101, 102
 Devolution, 2:761, 783
 Dewey, John, 2:619–620
 Dialectical idealism, 2:649
 Dialectical materialism, 2:651
 Diamond, Larry, 1:222, 259, 261, 262–263, 382
 Diamond, Martin, 1:63, 64, 66
 Diesing, Paul, 1:35
 Diffusion effect, 1:280
 Dillon’s Rule, 2:701
 Diminished subtype of democracy, 1:260, 262
 Direct elections, 1:159
 Direct lobbying, 2:756
 Disciplined-configurative case study, 1:295
Discourse on the Origins of Inequality (Rousseau), 1:592
 Discursive theory on nonstate actors, 1:417
 Distribution scarcity, 1:243
 District elections, 2:704
 District magnitude, 1:160
 Divided party government, 2:727
 Divine right, 1:210, 2:588
 DNA, 2:637
 Doha Round, 1:412
 Dominant party regime, 1:251
 Double institution, 1:78
 Downs, Anthony, 1:37, 40, 2:815
 Druckman, J. N., 2:713
 Drury, Shadia, 1:65
 Du Bois, W. E. B., 2:836
 Dual federalism, 2:761
 Duchacek, I. D., 1:170, 171
 Dummy variables, 2:486–487
 Duration analysis, 2:502–503
 Durkheim, Émile, 1:75, 194, 211
 Duverger’s Law, 1:161–162, 2:452, 453

 Earth Summit, 1:443, 445
 Easton, David, 1:6, 9–10, 30, 71, 72–74, 93, 174, 2:463, 464–465, 786
 Ecofeminists, 1:444
 Ecological footprint, 1:135
 Ecological modernization, 1:444
 Ecological parties, 1:151
 Ecological reserve, 1:135
 Ecology, international, 1:309
 Economic determinism, 2:650–652, 665
 Economic diversification, 1:399
 Economic equality, 1:7
 Economic law, international, 1:438
 Economic man, 1:35
 Economic model of rational choice
 administrative man concept in, 1:35–36
 criticism of, 1:39
 economic man concept in, 1:35
 Economic nationalism/realism, 1:408–409
 Economic policy, presidency and, 2:732
 Economics
 bureaucracy and, 2:748–749
 as cause of ethnic conflict, 1:145
 effects of terrorism on, 1:122
 See also **International political economy and trade**
 Economism, 2:662–663
 Economy
 planned, 1:254–255
 transfer, 1:254
 Ecumenism, 1:9–10
 Education, public, 1:174, 2:599
 Egalitarianism, 1:15–16
 E-government, 2:783
 Egypt
 elections in, 1:161, 250
 Mediterranean Action Plan, 1:427
 social movements in, 1:231
 Eisenhower, Dwight D., 2:728, 800
 Election rules, 1:154
 Election theory, 1:161–163
 Electoral authoritarianism, 1:260–261, 262
 Electoral College, 1:159–160, 178
 Electoral engineering, 1:162
 Electoral formula, 1:160
 Electoral reform, 1:164
Electoral system in comparative perspective, 1:159–167
 additional member system, 1:160–161
 applications and empirical evidence, 1:163–165
 closed list party system, 1:160
 direct/indirect elections, 1:159
 district magnitude and, 1:160
 election theory and, 1:161–163
 Electoral College, 1:159–160
 electoral formula and, 1:160
 electoral thresholds and, 1:160
 first-past-the-post system and, 1:160–161
 future research issues, 1:165–166
 mixed-member system, 1:161
 multimember districts, 1:160
 open list party system, 1:160
 party-list proportional system, 1:160, 161
 policy implications, 1:165
 proportional representation, 1:160
 single transferrable vote and, 1:160

single-member districts, **1:160, 161, 163**
 theory on, **1:161–163**
 Westminster-style parliamentary system, **1:160**
See also **Comparative political parties: Systems and organizations**
 Electoral thresholds, **1:160**
 Elite pacts argument, **1:279**
 Elite parties, **1:153**
 Elite political instability, **1:126**
 Elite theory, **2:755**
 Elitism, rejection of, **2:678**
 Embedded democracy, **1:260**
 Endogenously timed elections, **1:82, 181**
 Engels, Freidrich, **1:409–410, 2:649, 650, 653, 655–656, 664–665**
See also **Marxism**
 England. *See* Great Britain
 English Bill of Rights, **2:685, 686**
 Enlightenment, **1:210–211, 2:686–687**
 on idealism and liberalism, **1:320–321**
 origins of, **2:589**
 Enterprise associations, **2:622**
 Environment
 globalization effect on, **1:403**
 social democracy and, **2:662**
See also **Comparative environmental politics and conflict; International environmental politics**
 Environmental international relations, **1:307–308**
 Environmental law, international, **1:438**
 Epistemic community, **1:308, 416, 417–418, 427, 430(n1)**
 Equal opportunity, **2:616**
 Equality of outcome, **2:616**
 Equifinality problem, **1:288**
 Equilibria, **2:536**
 Erfurt Program, **2:656, 661**
 Esoteric, **2:558**
 Essentialist primordialism, **1:218**
Ethics (Aristotle), **2:557, 559**
Ethnic and identity politics, 1:217–224
 census elections and, **1:220**
 census outbidding and, **1:220**
 defining ethnic identity, **1:217–218**
 ethnic expressive voting, **1:221**
 ethnic parties and, **1:221–223**
 ethnic voting, **1:220–221**
 ethnic/national identity relationship, **1:219–220**
 expressive voting hypothesis and, **1:220**
 formation of ethnic identity, **1:218–219**
 LGBT persons and identity, **2:862**
 research issues, **1:220–223**
 resurgent identity, **1:219**

Ethnic conflict, 1:141–149
 clash of civilizations, **1:147–148**
 dynamics of, **1:145–146**
 escalation of, **1:146**
 ethnic groups worldwide and, **1:146–147**
 ethnic identity and, **1:141–143**
 ethnicity and, **1:141**
 media and, **1:145**
 origin/nature of, **1:143–146**
 proximate causes of, **1:144**
 underlying causes of, **1:144–145**
 Ethnic expressive voting, **1:221**
 Ethnic geography, **1:144**
 Ethnic groups
 defining, **1:141**
 voting behavior and, **1:220–121**
Ethnicity. See Race, ethnicity, and politics
 Ethnoclasses, **1:143**
 Ethnography, institutional, **1:19**
 Ethnonationalists, **1:143**
 Ethnoneutrality, **2:666**
 European Convention on Human Rights, **1:190**
 European Union, **1:172–173**
See also individual country
 Evans, Peter, **1:104–105**
 Evolutionary socialists, **2:654**
 Exclusive democracy, **1:260**
 Exoteric, **2:558**
 Expansionism, **1:305**
 Expansionist nationalism, **2:636**
 Expected utility theory (EUT), **1:379–380**
Experiments, 2:523–531
 applications, **2:528–530**
 experimental method, **2:523–525**
 field experiments on voter mobilization, **2:528–529**
 future research issues, **2:530**
 intervention comparison, **2:524–525**
 laboratory experiments on negative advertising, **2:528**
 laboratory experiments vs. field experiments, **2:526–527**
 limitations/need for replication, **2:529–530**
 natural experiments, **2:527–528**
 placebo effect, **2:525**
 policy implications, **2:530**
 random assignment, **2:524**
 randomized experiments vs. survey data, **2:525–526**
 survey-based experiments on racial attitudes, **2:529**
 theory on, **2:523–528**
 treatment and control groups, **2:524**
 uses or, **2:529**
 Exported revolution, **1:253**

- Expressive voting hypothesis, 1:220
 Extinction, of species, 1:440
 Extrasocietal system, 1:72–73
- Face-to-face interviews, 2:517
 Factual recall, 2:516
 Failed states
 civil wars and, 1:109, 110
 ethnic conflict and, 1:144
 Fairly uncomplicated experience argument, 1:279
 Faletto, Enzo, 1:104
 Falsificationism, 2:454–455
 Farmer, John D., 1:13
- Fascism and national socialism, 1:254, 2:639–647,**
 actual idealism and, 2:642
 biological determinism, 2:642–643
 empirical studies, 2:639–640
 fascist movement theory, 2:646
 futurism and, 2:642
 Germany and, 2:644
 historical context, 2:643–644
 Hitler and, 2:642, 643, 645–646
 ideological precursors, 2:640–643
 Italy and, 2:643–644
 leaders, 2:644–646
 Mussolini and, 2:642, 643, 644–645
 mysticism, 2:641–642
 national romanticism and, 2:640–641
 national socialism, 2:640
 nationalism, 2:640–641
 political spectrum studies, 2:640
 revolutionary syndicalism and, 2:641
 studies of fascism, 2:639–640
- Federal Court of Appeals, 2:741
 Federal Election Campaign Act, 2:801–802
 Federal system, 1:168–170, 173–174
 Federalism, 2:693–694, 695
 See also **Federalism, American**
Federalism, American, 2:760–768
 Article I, 2:763–765
 Article II, 2:765
 Article III, 2:765
 Article IV, 2:765–766
 Article V, 2:766
 Article VI, 2:766
 Article VII, 2:766
 Constitutional amendments, 2:766–767
 Constitutional Convention, 2:762–763
 dual federalism and, 2:761
 federalism in Constitution, 2:763–767
 initial plans for union, 2:762
 models of, 2:760–761
 nullification concept in, 2:760–761
 origins of, 2:761–763
 Preamble, 2:763
 Supreme Court cases, 2:767
Federalist, 2:606
- Feminists
 anarchist, 2:629
 ecofeminists, 1:444
 See also **International relations, feminist**
- Femocrats, 1:237
 Fenno, Richard, 2:717, 718
 Festinger, Leon, 1:55–56
 Feudalism, 1:278, 2:651
 Feuerbach, Ludwig, 2:664
 Fichte, Johann Gottlieb, 2:640
 Field experiments, 2:527, 528–529, 819
 Filmer, Robert, 2:590
 Finke, R., 2:850
 Fiorina, M., 2:817
 First Continental Congress, 2:688
 First-past-the-post party system, 1:154, 160–161
 Foreign aid, 1:87
 as rent, 1:396–397
 Foreign direct investment (FDI), 1:410
- Foreign policy, American, 2:822–830**
 bureaucracy and, 2:826–827
 Congress and, 2:825–826
 governmental politics model, 2:826, 827
 international affairs, U.S. role in, 2:823–824
 organizational behavior model, 2:826, 827
 president and, 2:824–825
 public opinion, 2:827–829
 rational actor model of, 2:826, 827
- Foreign policy analysis (FPA), 1:336–343**
 applications/empirical evidence, 1:338–141
 comparative foreign policy, 1:338
 constructivism on, 1:337
 contemporary approaches to, 1:338
 defining foreign policy, 1:336–337
 feminist theory on, 1:337
 foundational texts of, 1:338
 future research issues, 1:341–342
 group as policy decision maker, 1:339–340
 individual as policy decision maker, 1:339
 international relations theory, 1:337–339
 introduction, 1:336–137
 liberalism, 1:337
 Marxist theory on, 1:337–338
 policy implications, 1:341
 presidency and, 2:733
 realism, 1:337
 society as policy decision maker, 1:339

- Foreign policy entrepreneurs, 2:825
 Forensic approach, 1:252
 Forgotten Man, 2:601–662
 Forgotten Woman, 2:601, 602
 Formal equality, 2:616
 Formal functional analysis, 1:75
 Formal modeling, 1:285, 286–287, 290
Formal theory and spatial modeling, 2:532–540
 defining formal theory, 2:532–533
 empirical methods and, 2:534
 future research issues, 2:538–539
 probabilistic theory and, 2:535
 prospect theory and, 2:535–536
 public choice and, 2:536–537
 quantitative methods and, 2:533–534
 rational choice, 2:534–536
 solving and testing formal models, 2:536
 spatial theory of voting, 2:537–538
 utility maximization and, 1:57, 2:534–535
 Formateur party, 1:180
 Foucault, Michel, 1:15–16
 Fourier, Charles, 2:655, 664
 Fragmentation of political parties, 1:155
 Framing approach, 1:228
 Framing effects, 2:712, 713
 France
 colonialism and, 1:127
 decentralized education in, 1:174
 devolution revolution and, 1:171
 nationalization in, 2:660
 Frank, Andre G., 1:101, 102, 103
 Frankfort School, 1:252
 Franklin, Benjamin, 2:687, 688, 690, 744, 762
 Free market, 1:277
 religious, 1:214
 Free riders, 1:228, 2:758
 French and Indian War, 2:687, 762
 French Revolution, 2:593–594, 621
 Freud, Sigmund, 1:54, 55, 211
 Friedrich, C. J., 1:170
 Front party, 1:251
 Fukuyama, Francis, 1:323
 Functional cleavages, in political parties, 1:151
 Functional requisite, 1:78
 Functionalism. *See* **Systems theory and structural functionalism**
 Fund-raising, campaign, 2:797, 2:801–802
 Futurism, 2:642
 Gabon, rentier effects in, 1:398–399
 Gadhafi, Mu'ammarr Muhammad al-, 2:678–679
 Galileo, 2:588, 589
Game theory, 1:287, 288–289, 290, 2:541–549
 in American politics, 2:546–547
 in comparative politics, 2:547–548
 definitions/terminology, 2:542–545
 deterrence theory and, 1:372–373
 extensions of, 2:544
 extensive form of, 2:544–545
 game tree in, 2:544–545
 history of, 2:546
 in international relations, 2:547
 Nash equilibrium and, 2:544
 normal form of, 2:543–544
 players and preferences, 2:542–543
 in political science, 2:546–548
 positive-sum game, 1:407–408
 prisoner's-dilemma game, 1:290, 2:543–544
 rational choice-driven, 1:285
 representation and solution concepts, 2:543–545
 stationary subgame perfect Nash equilibrium, 2:546–547
 Gandhi, Mahatma, 2:564
 Garvey, Marcus, 2:836–837
 GATT, 1:409, 1:412
 Gauss-Markov assumptions, 2:484–486
 Gay liberation, 2:857–858
 Gay marriage, 2:781
 Geddes, Barbara, 1:96
 Gellner, Ernest, 2:634, 637
Gender and politics, 1:233–241
 evolution of study of, 1:233–234
 feminist comparative policy, 1:236
 policy formation and implementation, 1:238
 policy transfer issues, 1:239–240
 political representation and, 1:234–236
 practical implications of, 1:239
 quotas and, 1:239
 research strategies, 1:234
 state feminism, 1:237–238
 welfare state and, 1:236–237
 women's movements, 1:230, 238–239
Gender and politics in the U.S., 2:840–846
 campaigns and, 2:800, 843
 courts and, 2:844, 845
 descriptive representation and, 2:841–842
 electoral campaigns and, 2:843
 future research issues, 2:845
 legislatures and, 2:843–844, 845
 participation, 2:841
 personal attitudes, 2:841–842

- political actors, 2:840–841
- political candidates, 2:841–843
- political institutions, 2:843–845
- political parties, 2:843
- presidency and, 2:845
- recruitment, 2:845
- voter evaluation of candidates, 2:842–843
- Gender mainstreaming, 1:237
- General good, 2:598
- General systems theory, 1:71
- General will, 2:634
- Geneva Conventions, 1:437–438
- Gentile, Emilio, 2:639–640
- Gentile, Giovanni, 2:642
- Geoculture, 1:331
- George III (England), 2:688, 689, 772
- Georgia, Republic of, 1:145
- Germany
 - fascism in, 2:644
 - as federal system, 1:169
 - social democracy in, 2:655–658, 660, 661
- Ghetto, 2:834
- Giddens, Anthony, 1:12, 13, 16–17, 19
- Gilpen, Robert, 1:100, 426
- Gingrich, Newt, 2:774
- Global citizenship, 1:404
- Global governance theory, 1:323–324
- Global security, 1:305–306
- Global social movements, 1:404
- Global trade. *See* **International political economy and trade**
- Globalism vs. globalization, 1:403
- Globalization. *See* **Complex interdependence and globalization**
- Gobineau, Arthur de, 2:642
- Godesberg Program, 2:600
- Goldman, Emma, 2:626–627
- Goodnow, Frank, 1:5, 2:459–460
- Gore, Al, 1:182, 354
- Gorgias* (Plato), 2:558
- Government formation, 1:155–156
- Governmental politics model (GPM), 2:826, 827
- Grand narrative of modernity, 1:13
- Grassroots lobbying, 2:756
- Grassroots parties, 2:771–772
- Great Britain
 - colonialism and, 1:127
 - devolution revolution and, 1:171
 - free trade and, 1:409
 - National Health Service in, 2:659
 - nationalization in, 2:660
 - party capability thesis and, 1:190
 - representation in, 1:181
- Great Cultural Revolution, 2:678
- Great Depression, 1:409, 2:761
- Great Leap Forward, 2:670, 678
- Great Purge, 2:668
- Great Society, 2:761
- Green, D. P., 1:40
- Green movement, 1:227
- Greenhouse gas (GHG) emissions, 1:440, 445
- Greenpeace, 1:445
- Grievances approach, 1:228
- Griswold v. Connecticut*, 2:693
- Groupthink, 1:56
- Gruber, L., 1:426–427
- Guardian coup, 1:126
- Gunther, Richard, 1:222
- Gurr, T. R., 1:111
- Habermas, Jürgen, 2:610–612
- Hall, P. A., 1:25, 26
- Hamas, 1:116
- Hancock, John, 2:690–691
- Hate crimes, 1:116
- Hate speech, 1:145
- Hayes, Rutherford B., 2:773
- Heckscher–Ohlin theory, 1:408
- Hegel, Georg, 2:664
- Hegemony, 1:315, 362–363, 409, 410–411
- Heilbroner, Robert, 1:442–443
- Held, David, 1:267
- Herder, Johann Gottlieb von, 2:634
- Heresthetics, 1:38
- Hermann, Margaret, 1:339, 356
- Herrnson, Paul, 2:798, 799
- Heuristic case study, 1:295
- Heywood, Andrew, 2:634
- Hezbollah, 1:116
- Historic materialism, 1:100, 2:665
- Historical institutionalism, 1:26, 252, 2:736–737
- Historical structuralism, 1:409–411
- Historicism, 2:558
- Hitler, Adolf, 1:133, 2:645–646, 658
- HIV/AIDS, 1:237, 2:861–862
- Hobbes, Thomas, 1:368, 2:590, 617, 686
- Hobsbawm, Eric, 1:219, 2:634–635
- Holism, 1:29–30
- Home rule, 2:702
- Homer-Dixon, Thomas, 1:242, 243, 246, 308
- Homo Economicus*, 1:52, 56, 58
- Homo Psychologicus*, 1:52, 57, 58
- Homophile, 2:857

Homosexuality. *See* **LGBT issues and the queer approach**

Honduras, military coup in, **1:130**

Hooker, Richard, **2:585**

Hoover, Herbert, **1:356, 2:773–774**

Horowitz, David, **1:111, 113, 130**

Horowitz, Donald, **1:218**

Hotman, Francis, **2:589**

Human Development Programme (U.N.), **1:446**

Human security, **1:349**

Humanitarian crises, **1:309, 323**

Hume, David, **2:454, 2:593**

Hunter, Floyd, **2:689**

Huntington, Samuel, **1:126, 127, 129, 147–148, 206, 212, 255, 276**

Hurricane Katrina, **1:136–137**

Hussein, Saddam, **2:824**

See also Iraq

Hybrid regimes, **1:180, 261–262**

Hyman, Herbert, **2:806**

Hyperpluralism, **2:700**

Hypodermic model of media effects, **2:711**

Ibn Bajjah, **2:573**

Ibn Khaldun, **2:573–574**

Ibn Rushd, **2:573**

Ibn Sina, **2:573**

Ibn Tufayl, **2:573**

Idealism

actual, **2:642**

dialectical, **2:649**

See also **Idealism and liberalism**

Idealism and liberalism, 1:319–326

democratic peace theory, **1:320, 325**

Enlightenment thought on, **1:320–321**

institutional liberalism, **1:324–325**

intellectual roots of, **1:320–322**

interdependence liberalism, **1:324**

post–World War I, **1:321–322**

post–World War II, **1:322–323**

revival after cold war, **1:323**

sociological liberalism, **1:323–324**

See also **Liberals, early moderns and classical**

Identity

ethnic, **1:141–143**

See also **Ethnic and identity politics**

Ideological distance, **1:154–155**

Ideological purity, **2:678**

Ideology

authoritarianism and, **1:255**

See also **Idealism and liberalism**

Imagined communities, **1:219–220, 2:635**

Immiseration theory, **2:656–657**

Impact and outcome studies, **2:792**

Imperialism, **1:100–101, 305, 328, 366, 410, 2:562, 642**

Import-substitution industrialization (ISI), **1:104, 333–334, 420**

Impossibility theorem, **1:37, 273, 2:608**

Incomes policy, **2:661**

Incrementalism, **1:34, 2:475–476, 499, 659, 732, 745, 790–791**

India

as federal system, **1:169, 172**

high courts in, **1:190**

representation in, **1:181**

See also **Asian political thought**

Indigenous peoples, **1:143**

Indigenous takeover, **1:253–254**

Indirect elections, **1:159**

Individual, within-case study, **1:294**

Individualism, **1:29, 2:602**

Individualist-civic nationalism, **1:220**

Individualistic political culture, **2:756–757**

Individualized justice vs. bureaucratic efficiency, **1:13, 19**

Indonesia

natural resources and, **1:85**

rentier effects in, **1:398–399**

Inductive approach, **1:284, 290**

Inductive fallacy, **1:295**

Industrial bourgeoisie, **2:652**

Industrialization, import-substitution and, **1:104**

Inequality regime, **1:236**

Influentials, **1:7**

Inglehart, Ronald, **1:212**

Inquisitorial judicial model, **1:190**

Inside lobbying, **2:756**

Institutes of the Christian Religion (Calvin), **2:584, 588**

Institutional ethnography, **1:19**

Institutional interest groups, **2:753–754**

Institutional liberalism, **1:324–325**

Institutional rational choice, **2:793**

Institutional trust. *See* Political trust

Institutionalism

historical, **1:26, 252, 2:736–737**

liberal, **1:409**

neoliberal, **1:425**

on religion, **1:215**

See also **Neoinstitutionalism**

Institutionalization, of balance of power, **1:363–364**

Institutionalized semiauthoritarianism, **1:263–264**

Institutions, defining, **1:78**

Instructed delegate, **2:718**

Instrumentalism, **1:142, 218–219, 2:635**

Interdependence. *See* **Complex interdependence and globalization**

- Interdependence liberalism, 1:324
- Interdeterminacy, 1:286
- Interest groups, 2:718, 722
See also **Interest groups and pluralism**
- Interest groups and pluralism, 2:752–759**
 free rider, 2:758
 interest groups formation, 2:757
 lobbyists, 2:756–757
 pluralism model, 2:754–755
 political action committees, 2:757
 political culture and, 2:757–758
 public/private, 2:754
 purpose of interest groups, 2:755–756
 types of interest groups, 2:753–754
- Intergovernmental network, 1:324
- International Committee of the Red Cross (ICRC), 1:438
- International economic law, 1:438
- International environmental law, 1:438
- International environmental politics, 1:440–448**
 deep ecology and, 1:443–444
 ecofeminists and, 1:444
 ecological modernization, 1:444
 liberal international relations and, 1:441–442
 mainstream theories of, 1:441–442
 measuring success/failure of, 1:444–446
 neo-Malthusian/anti-Malthusian debate, 1:442–443
 policy implications and future directions, 1:446–447
 public choice theory and, 1:441
 radical theories of, 1:443–444
 transnational NGOs and, 1:442
- International governmental organizations (IGOs), 1:415, 418
- International human rights, 1:349–350
- International institutions, defining, 1:423
- International law, 1:431–39**
 customary law and, 1:434
 general principles of, 1:434
 Geneva Conventions and, 1:437–438
 history of, 1:432–433
 international economic law and, 1:438
 international environmental law and, 1:438
 international relations and, 1:431–432
 jus ad bellum, 1:437
 jus in bello, 1:437
 Kyoto Protocol, 1:438, 446, 2:824
 law of armed conflict, 1:437–438
 legally nonbinding rules (soft law) and, 1:434
 lex maritima, 1:438
 lex mercatoria, 1:438
 national law and, 1:432
 nonstate actors and, 1:436–437
 states and, 1:435–436
 treaties and, 1:433–434
- International Monetary Fund (IMF), 1:305, 329, 409, 429, 438
- International nongovernmental organizations (INGOs), 1:415, 442
- International organizations and regimes, 1:422–430**
 constructivism on, 1:427
 delegation and agency, 1:429
 international institutions, defining, 1:423
 international organizations, defining, 1:423–424
 international regimes, defining, 1:423
 neoliberal institutionalism, 1:424–426
 rational design and delegation, 1:428–429
 realism on, 1:426–427
- International political economy and trade, 1:407–413**
 economic nationalism/realism and, 1:408–409
 global trade growth, 1:411
 historical structuralism, 1:409–411
 international trade theory, 1:407–411
 liberalism on, 1:407–408
 major developments in, 1:411–412
 mercantilism and, 1:408–409
 multilateral institutions and, 1:411–412
 neomercantilism and, 1:408, 409
 protectionism and, 1:409
 regionalism impact on, 1:411
- International public goods, 1:425
- International regimes, 1:305, 324, 423
- International relations, 1:431–432
 example of case study on, 1:298–299
 game theory in, 2:547
 international law and, 1:431–432
 Straussians on, 1:64–65
See also **International relations, feminist;**
International relations, history of; Nonstate actors in international relations
- International relations, feminist, 1:306–307, 344–352**
 evolution of, 1:346
 future research issues, 1:350–351
 international human rights and, 1:349–350
 methodology, 1:346–347
 policy implications, 1:350
 sovereignty/state in, 1:347–348
 traditional IR vs., 1:346
 war/militarism/security in, 1:348–349
- International relations, history of, 1:303–310**
 constructivist international relations, 1:307
 cyber crime, 1:309
 development strategies and humanitarian crises, 1:309
 ecological challenges, 1:309
 economic interdependence and, 1:305
 environmental international relations, 1:307–308

- feminist international relations (*see* International relations, feminist)
- future research issues, 1:308–309
- global security and, 1:305–306
- historical/theoretical developments, 1:304–308
- liberal beginnings of, 1:304
- Marxist alternatives to, 1:304–305
- new security threats, 1:308–309
- realist critics of, 1:304
- International relations liberalism on nonstate actors, 1:416
- International trade theory, 1:407–411
- International Whaling Commission (IWC), 1:445
- Interpretivism, 1:252, 2:455–456, 467
- Intersectionality, 2:629
- Interstate wars
 - civil wars *vs.*, 1:107–108
 - See also* **Rivalry, conflict, and interstate war**
- Intersubjective, 2:471–472
- Interviews, face-to-face, 2:517
- Intrasocietal system, 1:72, 73
- Invisible hand theory, 1:30
- Inward-looking social capital, 1:195
- Iran
 - democracy and, 1:85
 - elections in, 1:250
 - Islam and, 1:143, 2:571, 574–575
 - language issues in, 1:174
 - nuclear weapons in, 1:352, 354, 378
 - Operation Ajax and, 1:127
 - rentier effects in, 1:398–399
 - revolution in, 1:252, 253
 - social movements in, 1:231
- Iraq
 - communism in, 2:674
 - coups and, 1:129
 - democracy and, 1:87, 214
 - elections in, 1:161, 162, 164
 - ethnic conflict in, 1:143
 - invasion of Kuwait, 1:118
 - language issues in, 1:174
 - military forces of, 1:129
 - nationalism in, 2:637
 - parliamentary government in, 1:181
 - peace movements and, 1:230
 - post-Saddam elections, 1:164
 - religious conflict in, 1:215
 - rivalry formation policy in, 1:382
 - socialism in, 2:677
 - Sunni insurgency in, 1:121–122
 - war, 2003, 1:325, 339, 341, 353–354, 382
 - war on terrorism and, 2:824
 - weapons of mass destruction and, 1:65, 353
- Irish Americans, 2:835
- Iron Guard, 2:646
- Iron law of wages, 2:650
- Iron triangle, 2:730–731
- Ishiyama, J., 2:492
- Islamic political thought, 2:568–576**
 - al-Farabi, 2:570, 572
 - classical philosophers, 2:572–574
 - Ibn Bajjah, 2:573
 - Ibn Khaldun, 2:573–574
 - Ibn Rushd, 2:573
 - Ibn Sina, 2:573
 - Ibn Tufayl, 2:573
 - imans, influence on, 2:569, 570–571
 - Islamism, 2:574
 - Kharijites school, 2:571–572
 - modern, 2:574–575
 - Muhammad and, 2:569, 571
 - premodern Islamic political thought, 2:569–574
 - Qur'an, 2:570
 - reformist fundamentalism, 2:574
 - Shari'a, 2:571
 - Shi'a school of, 2:570–571
 - Sunni school of, 2:570
 - traditionalist fundamentalism, 2:574
- Islamism, 2:574
- Isomorphism, 1:402
- Israel
 - Islamic political thought and, 2:574
 - nationalism in, 2:635
 - nuclear weapons in, 1:378
 - political culture in, 1:341
 - Israeli–Palestinian conflict, 1:121, 144
- Italy, 1:171, 2:643–645
 - See also* **Fascism and national socialism**
- Jacobs, Jane, 2:700
- Jaffa, Harry, 1:64, 65
- Japan
 - fascist movement in, 2:646
 - See also* **Asian political thought**
- Jefferson, Thomas, 2:686, 687, 688–689, 690, 727, 772, 852
- Jervis, Robert, 1:369, 377–378, 380
- Jews, 1:117, 143, 2:643, 645, 646
- John of Salisbury, 2:580–581
- Johnson, Lyndon B., 1:356, 2:706, 761, 774
- Judeo-Christian teachings, influence on American politics, 2:687

- Judicial decision making. *See* **Comparative judicial politics; Judicial politics, American**
- Judicial politics, American, 2:735–742**
 attitudinal model on, 2:735–736, 738
 historical institutionalism on, 2:736–737
 judicial motivation, 2:741
 judicial opinions, importance of, 2:741
 measurement of attitudes, 2:741
 policy implications, 2:740
 political science model on case outcomes, 2:739
 precedent, 2:739
 role conceptions, 2:738
 social diversity of judges, 2:737–738
 strategic model of judicial behavior, 2:736, 739
 theory on, 2:735–737
 voting patterns, 2:738
- Judicial Research Initiative, 2:741
- Judicial systems, state, 2:782–783
- Judiciary. *See* **Judicial politics, American**
- Juries, 1:190–91
- Jus ad bellum, 1:437
- Jus in bello, 1:437
- Kant, Immanuel, 1:320, 387, 2:589, 664, 828
- Kantian revolution, 2:587, 594
- Kardinger's Basic Personality Type, 2:806–807
- Karl, Terry Lynn, 1:85
- Kasza, Greg, 1:9–10
- Kautsky, J. H., 2:673–674, 676
- Kautsky, Karl, 2:655, 657
- Keech, Marian, 1:55–56
- Keman, H., 1:156
- Kennedy, John F., 1:354
- Kennedy, Robert, 1:354
- Keohane, Robert, 1:314, 322–323, 324, 346, 2:457
- Kerala, communism in, 2:674
- Keynes, John M., 2:660
- Kharijites school, 2:571–572
- King, G., 2:457
- Kinsey, Alfred, 2:856
- Kinship primordialism, 1:218
- Kissinger, Henry, 1:313, 382
- Knowledge, value of, 1:11, 19
- Know-Nothings, 2:835
- Kornegger, Peggy, 2:629
- Kosovo, 1:145, 433
- Kristallnacht (Crystal Night), 2:646
- Kropotkin, Peter, 2:626
- Ku Klux Klan (KKK), 1:118, 2:712
- Kuhn, Thomas, 1:14, 2:455–456, 457
- Kurds, 1:108, 143, 174
- Kyoto Protocol, 1:438, 446, 2:824
- Labor incorporation study, 1:289
- Laboratory experiments
 field experiments vs., 2:526–527
 on negative advertising, 2:528
- Lagged time-series data, 1:289–290
- Lahiri, S., 1:134–135
- Lakatos, Imre, 2:456–457
- Language, construction of reality and, 1:15, 19
- Laos, communism in, 2:674
- Large-*N* studies, 1:288
- Lasswell, Harold, 1:54
- Latent content, 2:492
- Latin America
 democratic breakdown in, 1:279
 social movements in, 1:231
See also individual country
- Law
 national law, 1:432, 2:694
 neoinstitutionalism and, 1:23
 religious system of, 1:187, 2:571
 soft law, 1:434
See also International law
- Law of armed conflict, 1:437–438
- Law of capital accumulation, 1:100
- Law of disproportionality, 1:100
- Leadership and decision making, 1:353–360**
 future research issues, 1:359
 importance of decision-making process, 1:354–355
 importance of studying political leadership, 1:353–334
 multivariate approaches to, 1:358–359
 personality and politics, 1:356
 political leaders as decision unit, 1:355–359
 psychoanalytic studies of leadership, 1:355–356
 single personality variables, 1:356–358
- Leadership Trait Analysis (LTA), 1:339
- League of Nations, 1:321, 377, 433
- Least developed countries (LDCs). *See* **Dependency and development**
- Least-likely case study, 1:295
- Least squares dummy variable (LSDV), 2:502
- Lebanon, 1:116
- Lee Kwan Yew, 2:565
- Legal rationality, 1:35
- Legally nonbinding rules (soft law), 1:434
- Legislatures, gender and, 2:843–844, 845
- Legitimate authority, defining, 1:249–250
- LeMay, Michael, 2:833–834, 836
- Lemon vs. Kurtzman*, 2:852
- Lenin, Vladimir, 1:100–101, 305, 410, 2:666–668

Leninism, communism, Stalinism, and Maoism, 2:664–672

communist ideology in 21st-century politics, 2:670–671

Leninism, 2:666–668

Maoism, 2:669–670

Marx–Engels theory of communism, 2:664–666

Stalinism, 2:668–669

Lesbian feminism, 2:857

Leviathan (Hobbes), 2:590, 617, 686

Lex maritima, 1:438

Lex mercatoria, 1:438

LGBT issues and the queer approach, 2:856–864

bisexual movement, 2:858–859

education/representation and, 2:860–861

gay liberation and, 2:857–858

health and, 2:861–862

HIV/AIDS, 1:237, 2:858, 861–862

identity politics, critique of, 2:862

legal change, 2:861

lesbian and gay movements in U.S., 2:856–858

lesbian feminism and, 2:857

LGBT politics, main issues in, 2:860–862

nationalism and, 2:863

neoimperialism and, 2:863

neoliberalism, critique of, 2:862–863

politics of visibility, critique of, 2:862

postwar period, 2:856–857

queer approaches, 2:862–863

same-sex partnerships, 2:861

terrorism and, 2:863

transsexual/transgender movement, 2:859–860

Liberal institutionalism, 1:409

Liberal international relations, environmental politics and, 1:441–442

Liberal minimalist theory, 2:608–610

Liberal nationalism, 2:636

Liberal parties, 1:151

Liberalism

emergence of, 1:592

international political economy and, 1:407–108

international relations and, 1:304

nonstate actors and, 1:416

See also **Idealism and liberalism; Liberalism, conservatism, and libertarianism, modern; Liberals, early moderns and classical; Neoclassical liberals**

Liberalism, conservatism, and libertarianism, modern, 2:615–624

Burke on, 2:621

conservatism, 2:620–622

equal opportunity, 2:616

equality of outcome, 2:616

formal equality, 2:616

human nature and political values, 2:615–616

libertarian theories of politics, 2:618–619

modern liberal theories of politics, 2:619–620

modern liberalism and libertarianism, 2:616–620

negative liberty, 2:616, 618, 619–620

origins of liberal tradition, 2:617–618

perennial questions and future prospects, 2:622–623

positive liberty, 2:616

redistribution concept, 2:616

Liberals, early moderns and classical, 2:587–595

Copernicus, 2:588–589

Enlightenment, 2:589

French Revolution, 2:593–594

Hobbes, 2:590

Kantian revolution, 2:594

Locke, 2:590–592

Protestant Reformation, 2:587–588

Rousseau, 2:592–593

royal powers justification, 2:589–590

social contract theory, 2:589–593, 590, 593

See also **Idealism and liberalism; Neoclassical liberals**

Liberation theology, 1:211, 213

Libertarian theories of politics, 2:618–619

Libertarianism, 2:616–620

Libya, socialism in, 2:678–679

Lieber, Francis, 1:4

Lijphart, Arend, 1:171, 270–271, 284, 294

Limongi, F., 1:86

Lincoln, Abraham, 1:64, 65, 2:727, 761

Lindholm, Charles, 1:34–35, 2:607, 745

Lippmann, Walter, 2:711

Lipset, S. M., 1:276, 277

Litt, Edgar, 2:833, 834–835

Local governments

county, 2:783–784

municipal, 2:784

Locke, John, 1:210–211, 320, 2:590–592, 617–618, 686

Logical positivism, 2:461

Lone countries, 1:147

Longdregan, John, 1:129

Longitudinal analysis, 2:497–505

cross-sectional vs., 2:498–499

current approaches, 2:500

data considerations, 2:500–501

duration analysis, 2:502–503

future research directions, 2:503–504

historical developments, 2:499–500

models, 2:501–502

- statistical programs, **2:503**
- survival analysis, **2:502–503**
- Long-Range Transboundary Air Pollution (LRTAP), **1:445**
- Long-term decrease in profits, **1:100**
- Lowell, Lawrence, **2:460**
- Luciani, G., **1:335–336**
- Luther, Martin, **1:210, 2:583–584, 587–588**
- Luttwak, Edward, **1:125–126, 126**
- Lyotard, Jean-François, **1:12, 13, 14, 19**

- MacDonald, S. E., **2:815**
- Macedonia, nationalism in, **2:637**
- Machiavelli, Niccoló, **1:210**
- Macridis, Roy C., **1:24**
- Madison, James, **1:196, 2:606, 692–693, 696–697, 753**
- Magna Carta, **2:685**
- Maistre, Joseph de, **2:621**
- Majoritarian (Westminster) democracy, **1:270–271**
- Majority rule, **1:591**
- Male breadwinner model, **1:236–237**
- Malthus, Thomas, **1:133, 242**
- Malthusianism, **1:442–443**
- Manifest content, **2:492**
- Many variables, small-*N* problem, **1:286**
- Mao Zedong, **1:252, 2:565**
- Maoism, **2:669–670, 677–678**
- Marbury vs. Madison*, **2:721, 739**
- March, James, **1:35–36**
- March, James G., **1:23, 24**
- Marginalized peoples, individual realities of, **1:19–20**
- Marinetti, Filippo Tommaso, **2:642**
- Market forces, neoclassical liberals on, **2:600**
- Market socialism, **2:678**
- Market theory, **2:850–852**
- Marketplace of ideas, **2:708**
- Marshall, T. H., **2:659**
- Marsilio de Padua, **2:583**
- Marx, Karl, **1:99–100, 211, 409–410, 2:560–561, 653–654, 655**
- See also **Marxism**
- Marxism, 2:648–654**
 - classical, **2:664–666**
 - collapse theory as pillar of, **2:652, 656**
 - dependency/development and, **1:99–100, 101, 102, 328**
 - dialectical materialism and, **2:650, 651**
 - economic determinism and, **2:650–652**
 - historical materialism and, **1:100**
 - human nature concept in, **2:650**
 - immiseration theory as pillar of, **2:656–657**
 - international relations and, **1:304–305**
 - labor theory of value in, **2:650**
 - means of production and, **2:651**
 - nationalism and, **2:669**
 - notion of the state in, **1:91, 93**
 - overview of, **2:650**
 - professional revolutionary role in, **2:653–654**
 - socialism/communism in, **2:653**
- Marxism-Leninism, **2:677, 679**
- Maslow, Abraham, **2:745**
- Mason, George, **2:696–697**
- Mason, T. D., **1:111–112**
- Mass parties, **1:153**
- Mass politics, **1:254**
- Mass spectacle, **1:255**
- Materialism
 - dialectical, **2:651**
 - historic, **1:100, 2:665**
 - Marxism and, **2:649**
- May Fourth Movement, **2:563**
- Mayhew, David, **2:717, 718**
- Mayor-council-chief administrative officer form, **2:703**
- McClosky, Herbert, **1:6–8**
- McCulloch v. Maryland*, **2:694**
- McDonaldization of criminal justice, **1:18**
- McKinley, William, **2:773**
- Means of production, **2:651**
- Mearsheimer, John, **1:382**
- Measurement issues, **1:287**
- Mechanical effects, in elections, **1:161–162**
- Mechanism model of testing, **1:288**
- Media and politics, 2:708–716**
 - agenda setting, **2:711–712**
 - basic models of media effects, **2:711–713**
 - bias in media, **2:709–710**
 - consolidation of media, **2:709**
 - democracy, media role in, **2:708–709**
 - ethnic conflict, **1:145**
 - framing effects, **2:712, 713**
 - going public, **2:710–711**
 - hypodermic model of media effects, **2:711**
 - media functions, **1:77**
 - minimal effects model, **2:711**
 - new media, **2:713–715**
 - new media effect on participation, **2:714–715**
 - ownership and regulation, **2:709**
 - political advertising, **2:710**
 - priming effects, **2:712, 713**
 - subtle media effects, **2:712–713**
- Media bias, **2:709–710**
- Median legislator, **1:162**
- Median voter theorem, **1:162**
- Mediterranean Action Plan, **1:427**
- Meiji Restoration, **2:563**

- Meinecke, Friedrich, **2:634**, **635**
 Mercantilism, **1:408–409**
 Merkl, Peter, **1:233**
 Merriam, Charles, **2:464**
 Merton, Robert, **1:75**
 Metanarrative, **1:14**
 Methodological individualism, **2:609**
 Methodology. *See* **Comparative methods**
 Mexican Americans, **2:836**
 Mexico, NAFTA and, **1:426–427**
 Micronarrative, **1:14**
 Middle class, democracy and, **1:82–83**, **278**
 Middle East
 communism in, **2:674**
 rentier effects in, **1:398–399**
 See also **Resources and rentierism, global politics of;**
 individual country
 Milgram, Stanley, **1:55**
 Mill, John Stuart, **1:320**, **2:595**, **618**
 Minimal effects model, **2:711**
 Minimal winning coalition, **1:156**
 Mitofsky–Waksberg (MW) method, **2:518**
 Mitterrand, François, **1:164**, **171**
 Mixed-member system, **1:161**
Models of democracy, 1:267–274
 ancients and democracy, **1:268–270**
 Arrow's impossibility theorem, **1:273**
 consensual democracy, **1:271**
 consociational democracy, **1:271**
 contemporary models of democracy, **1:270–273**
 delegative democracy, **1:271–272**
 deliberative democracy, **1:272**
 democratic autonomy, **1:272–273**
 majoritarian (Westminster) democracy, **1:270–271**
 participatory democracy, **1:269**
 polyarchy, **1:267–268**, **270**
 republicanism, **1:269–270**
 Moderate multipartism system, **1:155**
Modern democratic thought, 2:605–614
 agonistic pluralist theory, **2:611–612**
 deliberative polling technique, **2:612–613**
 deliberative theory, **2:610–611**
 democratic theory, **2:606–612**
 democratic thought, **2:606**
 democratic transitions, **2:612**
 future research issues, **2:613**
 liberal minimalist theory, **2:608–610**
 new pluralism, **2:613**
 participatory theory, **2:607–609**, **612**
 pluralist theory, **2:606–607**
 political science, **2:605–606**
 Modern world-systems (MWS) theory, **1:102–103**
 Modernist approach, to nationalism, **2:635–636**
 Modernity
 grand narrative of, **1:13**
 meaning of, **1:12**
 radicalized, **1:17**, **1:19**
 Modernization
 ecological, **1:444**
 as linear and nonreversible, **1:327–328**
 See also **Political development and modernization**
 Modernization effect, **1:86**
 Modernization hypothesis, on voter dealignment, **1:157**
 Mohamad, Mahathir, **1:254**
 Mol, Arthur, **1:447**
 Montesquieu, C., **1:92**, **210–211**, **2:692–693**, **763**
 Montreal Protocol on Substances That Deplete the Ozone Layer, **1:445**
 Moral hazard, **1:43**
 Moral relativism, **2:622**
 Moralistic political culture, **2:757**
 Morgenthau, Hans J., **1:312**, **313**, **362**, **365**, **377**
 Mortgage debt, toxic, **1:48–49**
 Most different systems, **1:284**
 Most-similar systems, **1:284**
 Most-likely case study, **1:295**
 Mouffe, C., **2:612**
 Muhammad, **2:569**, **571**
 Multilateral institutions, **1:411–412**
 Multimember districts, **1:160**
 Multinational corporations (MNCs), **1:330–331**, **334**,
 408, **411**, **412**, **416**, **417**, **418**
 Multiparty system, **1:154**
 Multiple-party system, **1:181**
 Multivariate approach, **1:358–359**
 Multivariate regression, **2:484**
 Municipal government, **2:784**
 Musharraf, Pervez, **1:127**, **128**
 Mussolini, Benito, **2:644–645**
 Mutually assured destruction (MAD), **1:305–306**
 Mysticism, **2:641–642**
 Myth of the general strike, **2:641**
 Naess, Arne, **1:443–444**
 Nasser, Gamal Abdel, **1:263**
 Nation, defining, **1:144**
 National Health Service (Britain), **2:659**
 National law, **1:432**, **2:694**
 National minorities, **1:143**
 National romanticism, **2:640–641**
 National socialism. *See* **Fascism and national socialism**
Nationalism, 2:633–638
 anticolonial, **2:636–637**
 collectivist-civic, **1:220**

- collectivist-ethnic, 1:220
- conservative, 2:636
- cuius regio, eius religio* concept in, 2:634
- cultural heterogeneity, 2:635
- defining, 1:144
- expansionist, 2:636
- imagined communities and, 2:635
- individualist-civic, 1:220
- instrumentalism and, 2:635
- LGBT issues and, 2:863
- liberal, 2:636
- Marxist view of, 2:669
- modernism and, 2:635–636
- perennialism and, 2:635
- political dimension of nations, 2:634–636
- political presentation of nationalism, 2:636–638
- primordialism and, 2:635
- state-nation concept and, 2:634, 635
- Nation-state
 - balance of power and, 1:362
 - defining, 1:93–94
- NATO, 1:354
- Natural experiments, 2:527
- Natural resources
 - curse of, 1:85–86
 - oil embargo and, 1:334
 - See also* **Resources and rentierism, global politics of**
- Natural rights, 1:62, 2:597–598, 602
- Natural things, 2:554
- Necessary law of nations, 1:320
- Negative campaign ads, 2:710, 801
- Negative liberty, 2:616, 618, 619–620
- Nehru, Jawaharal, 2:564
- Neoclassical liberals, 2:596–604**
 - on Forgotten Man, 2:601–602
 - on Forgotten Woman, 2:601, 602
 - on general good, 2:598
 - on individualism, 2:602
 - on market forces, 2:600
 - on minimal government, 2:601
 - on natural penalties, 2:601, 602
 - on natural rights, 2:602
 - on protection of natural rights, 2:597–598
 - on public assistance, 2:597–598
 - on public education, 2:599
 - on role of government, 2:597, 600–601
 - science and, 2:602–603
 - self-adjusting principle of, 2:598, 600
 - Spencer on, 2:596–600
 - Sumner on, 2:600–603
 - on war, 2:599
- Neocolonialism, dependency theory on, 1:329
- Neo-Gramscianism, 1:410–411
- Ne imperialism, LGBT issues and, 2:863
- Neoinstitutionalism, 1:22–28**
 - behavioral revolution as root of, 1:22–23
 - in comparative politics, 1:24
 - emergence of, 1:23–24
 - historical institutionalism, 1:26
 - law/legal analysis and, 1:23
 - levels of analysis, multiple, 1:24–25
 - political institution, defining, 1:24
 - on presidency, 2:727
 - rational choice institutionalism, 1:25
 - sociological institutionalism, 1:26
 - traditionalists and, 1:22–23
 - See also* Institutionalism
- Neoliberal institutionalism, 1:424–426
- Neoliberalism
 - economic interdependence and, 1:404
 - LGBT issues and, 2:862–863
- Neo-Malthusism, 1:133, 134, 136, 442–443
- Neomercantilism, 1:408, 409
- Neopatrimonial state, 1:94–95
- Neopluralism, 2:700
- Neorealism. *See* **Realism and neorealism**
- Network analysis, 1:285
- New Deal, 2:744, 761
- New Federalism, 2:761
- New Left, 1:151
- New media, 2:713–715
- New penology, 1:18
- New pluralism, 2:613
- New Right, 1:151
- New social movement, 1:227
- Newly industrialized countries, 1:104
- Newly industrializing country (NIC), 1:409
- News media, construction of reality by, 1:17
- Newton, Isaac, 2:589
- Nietzsche, Friedrich, 1:211, 2:594, 643
- Nigeria
 - nationalism in, 2:637
 - natural resources and, 1:85
 - religious conflict in, 1:215
 - rentier effects in, 1:398–399
- Nisbet, Robert, 2:621–622
- Nixon, Richard, 1:313, 1:356, 2:761, 2:774, 2:825
- No Child Left Behind Act, 2:701
- Nonassociational interest groups, 2:753
- Nondenominational churches, 2:850
- Nongovernmental organizations (NGOs), 1:414, 415, 418–419
- Nonprobability sampling, 2:519–520

Nonstate actors, 1:436–437

Nonstate actors in international relations, 1:414–421

- accountability of, 1:419
- applications and agency of, 1:417–419
- constructivism on, 1:416
- discursive theory on, 1:417
- future research issues, 1:420
- grassroots organizations, 1:418
- history and nomenclature, 1:414–416
- international governmental organizations, 1:415, 418
- international nongovernmental organizations, 1:415
- international relations liberalism on, 1:416
- legitimacy of, 1:419
- multinational corporations, 1:330, 334, 408, 411, 412, 416, 417, 418
- neutrality of, 1:419–420
- nongovernmental organizations, 1:414, 415, 418–419
- policy implications, 1:420
- rational choice and, 1:419
- realist theory on, 1:416
- regime theory on, 1:416
- religious organizations as, 1:418
- turbulence theory on, 1:416

Nonstate terrorism, 1:325

Nordlinger, Eric, 1:92, 93, 126

Normal science, 1:14

Norris, Pippa, 1:212

North American Treaty Organization (NATO), 1:428–429, 2:637

North Korea

- nuclear power in, 1:352
- nuclear weapons in, 1:352, 354
- socialism in, 2:679

North Vietnam, 2:674

Northern Ireland, 1:144, 171

Nozick, Robert, 2:619

Nuclear age, balance of power in, 1:365–366

Nuclear Non-Proliferation Treaty (NPT), 1:340–341

Nuclear war, 1:368–369

Nuclear weapons, 1:305–306, 352, 354, 378

Oakeshott, Michael, 2:622

Obama, Barak, 2:728, 761, 775, 837–838

Objective rationality, 1:35

Objective truth, 1:15, 16

Objectivism, postmodernism and, 1:12

Occupational niche, 2:834

Occupational queuing, 2:834

O'Donnell, Guillermo, 1:271–272

Of the Laws of the Ecclesiastical Polity (Hooker), 2:585

Officialism, 2:597, 600

Oil embargo, 1:334

Olsen, Johan P., 1:23, 24

Olson, Mancur, 1:38, 40, 44, 2:754, 758, 815, 816

On Liberty (Mill), 2:618

On the Revolution of the Celestial Spheres (Copernicus), 2:588

Ontological anarchism, 2:628–629

Open-ended questions, 2:515–516

Open list party system, 1:160

Open-system administration, 2:787

Operation Ajax, 1:127

Operational code analysis, 1:57

Operations research, 2:792

Ophuls, William, 1:442–443

Ordinal ballot, 1:160

Ordinary least squares (OLS) regression, 2:478

Organic nationalism, 2:635

Organization of Arab Petroleum Exporting Countries (OAPEC), 1:334

Organizational behavior model (OBM), 2:826, 827

Organizational process model, 2:790–791

Origins of Species (Darwin), 2:665–666

Orthodox variant of socialism, 2:677

Oslo Accords, 1:341

Ostrom, Elinor, 1:39, 40, 441, 2:793

Ostrom, Vincent, 1:39, 40

Output legitimacy, 1:261

Outside lobbying, 2:756

Outward-looking social capital, 1:195

Oversoul, 2:563

Owen, Robert, 2:655

Paine, Thomas, 2:593, 619

Pakistan, 1:127, 128

Palestine

Israeli–Palestinian conflict, 1:121, 144

nationalism in, 2:635

Pangle, T. L., 1:63, 66

Panopticism, 1:16, 129

Pape, R. A., 1:122–123

Paradox of plenty, 1:85

Pareto's efficiency, 2:536

Parliamentarism. *See* **Presidentialism versus parliamentarism**

Parliamentary supremacy, 1:189–190

Parochial political culture, 1:201, 2:756

Parsons, Talcott, 1:72, 75, 82

Participant political culture, 1:202, 2:756

Participation, political

gender affect on, 2:841

new media effect on, 2:714–715

theory on, 2:607–608

- Participatory democracy, 1:269
- Partisan campaigns, 2:799–800
- Partisan realignment, 1:156–157
- Party families, 1:151
- Party identification approach, 1:56, 2:770
- Party in central office, 1:152
- Party in government, 2:770
- Party in public office, 1:152–153
- Party in the electorate, 2:770
- Party on the ground, 1:152
- Party organizations, 1:152–154
- Party regime, 1:250–251
- Party system, 1:154–155
- See also* **Comparative political parties: Systems and organizations**
- Party-list proportional system, 1:160, 161
- Passive resistance, 2:584
- Pateman, C., 2:607–608
- Path dependence, 1:26
- Patriarcha* (Filmer), 2:590
- Patronage democracy, 1:221
- Patron–client relationship, 1:95–96
- Pattern variables, 1:72
- Peace of Westphalia, 1:432
- Perceptual factors, in ethnic conflict, 1:145
- Perennialism, 2:635
- Perestroika protest in political science, 1:9–10, 40
- Permanent campaign, 2:797
- Permanent revolution, 2:678
- Perry, Matthew, 2:563
- Personality studies, 1:54–55
- Perspective theory, 1:40
- Petite bourgeoisie, 2:652
- Petition of Rights, 2:685–686
- Philippines, high courts in, 1:190
- Philosophical anarchism, 2:630
- Pierce, Franklin, 2:772
- Pierson, P., 1:26
- Pipes, Richard, 2:668
- Placebo effect, 2:525
- Plato, 1:62, 65, 269, 2:553–554, 557, 558–559
- Plausibility probes, 1:295
- Plessey v. Ferguson*, 2:833
- Pluralism, 2:606–607
 - agonistic, 2:611–612
 - democratic, 1:254
 - hyperpluralism, 2:700
- See also* **Interest groups and pluralism**
- Poetic terrorism, 2:628
- Point predictions, 2:536
- Polanyi, K., 2:655
- Polarity, balance of power and, 1:362–363
- Polarized multipartism system, 1:155
- Policratus* (John of Salisbury), 2:580–581
- Policy framework approach, 1:221
- Policy implications, 1:165
- Policy paradox, 2:793
- Policy stream, 2:826
- Polis, 2:554
- Political Action Committees (PACs), 2:799–800
- Political advertising, 2:710
- Political and military coups, 1:124–132**
 - acute causes of, 1:126–127
 - breakthrough coup, 1:126
 - chronic causes of, 1:127
 - consequences of, 1:128
 - definition of coups d'état, 1:124–126
 - as elite political instability, 1:126
 - extralegal/illegal transfers of power in, 1:125
 - foreign policy responses to external, 1:130–131
 - future research issues, 1:131–132
 - guardian coup, 1:126
 - internal state policies on, 1:128–130
 - nature and causes of coups d'état, 1:124–127
 - policy implications of, 1:127–131
 - popular will and, 1:125–126
 - seesaw coup, 1:130
 - self-coup, 1:126
 - veto coup, 1:126
 - violence as noncritical in, 1:125
- Political attitude research, 1:55–56
- Political conditionality, 1:165
- Political consultants, 2:798–799
- Political culture, 1:201–208**
 - absolute monarchy, 1:202
 - civic culture and, 1:201–202
 - contemporary, 1:206
 - critical citizen in, 1:207
 - parochial, 1:201
 - participant, 1:202
 - political efficacy of, 1:203
 - political interest and knowledge in, 1:204–205
 - political participation in, 1:205–206
 - political trust and, 1:203–204
 - social capital and, 1:202–203
 - subject, 1:201–202
 - types of, 2:756–757
- Political development and modernization, 1:81–90**
 - criticisms of connection between, 1:86
 - dependency theory and, 1:84–85
 - economic growth and emergence of democracy, 1:84–85
 - future research issues, 1:88–89

- modernization, 1:82–83
- natural resource curse and, 1:85–86
- policy and promotion of democratization, 1:87–88
- political order in changing societies, 1:84
- selectorate theory and, 1:87, 88
- social mobilization and, 1:86
- survival and, 1:86
- Political equality, 1:7
- Political instability pathway, 1:127
- Political institution
 - defining, 1:24
 - See also* **Neoinstitutionalism**
- Political machines, 2:272, 702, 769
- Political man concept, 1:54
- Political opportunity structure approach, 1:227–228
- Political parties
 - gender and, 2:843
 - See also* **Comparative political parties: Systems and organizations; Political parties, American**
- Political parties, American, 2:769–778**
 - campaign financing, 2:776–777
 - components of, 2:770–772
 - defining, 2:769–770
 - functions of, 2:770
 - grassroots parties, 2:771–772
 - national committees, 2:771
 - organization of party, 2:771
 - party identification, 2:770
 - party in government, 2:770
 - party in the electorate, 2:770
 - political machines, 2:272, 769
 - state organizations, 2:771
 - third parties, 2:775–776
 - two-party system, 2:772–775
 - voters and, 2:770–771
- Political party formation, 1:150–152
- Political personality, 1:54
- Political psychology, 1:51–59**
 - central assumptions of, 1:51–53
 - coup studies and, 1:130
 - decision making and international politics, 1:56–57
 - emotion/neuroscience and, 1:57–58
 - methodology in, 1:53
 - personality studies, 1:54–55
 - roots of, 1:53–54
 - subfields of, 1:52–53
 - voting behavior/political attitude research, 1:55–56
- Political rationality, 1:35
- Political scarcity, 1:136
- Political science
 - game theory in, 2:546–548
 - modern democratic thought and, 2:605–606
 - Political science model on court case outcomes, 2:739
- Political socialization, 2:805–812**
 - agents of, 2:807–808
 - factors impacting, 2:806–807
 - how process works, 2:810
 - problems in study of, 2:810–811
 - recipients of, 2:808–809
 - significance of, 2:806
 - when socialization occurs, 2:809–810
- Political spectrum studies, 2:640
- Political stream, 2:826
- Political surplus, 1:102
- Political systems, founding of American, 2:685–698**
 - anti-federalist view, 2:695
 - Articles of Confederation, 2:689–690
 - basic features of Constitution, 2:692–694
 - British legal history, 2:685–686
 - checks and balances, 2:693
 - classical republicanism and, 2:687
 - colonies transformation into states, 2:689
 - Constitution, 2:690–697
 - Constitutional convention, 2:690–691
 - Constitutional development, influences on, 2:685–687
 - Declaration of Independence, 2:688–689
 - economic issues, 2:692
 - federalist system, 2:693–694
 - federalist view, 2:695
 - First Continental Congress, 2:688
 - Judeo-Christian teachings and, 2:687
 - liberal arts/Enlightenment and, 2:686–687
 - natural rights philosophers and, 2:686
 - new government and Bill of Rights, 2:696–697
 - ratifying Constitution, 2:694
 - representation issues, 2:691–692
 - salutary neglect influence on, 2:687–688
 - Second Continental Congress, 2:688, 689
 - separation of powers and, 2:692–693
 - slavery issues, 2:692
 - supremacy of national law, 2:694
- Political trust, 1:203–204
- Political violence
 - internal, vs. civil war, 1:108–110
 - See also* **Terrorism**
- Politicians' dilemma, 1:95–97
- Politics* (Aristotle), 2:557, 559
- Politics of dependence, 2:565–556
- Polyarchy, 1:267–268, 270, 2:606
- Poole, Keith, 1:129
- Popkin, S. L., 2:816
- Popper, Karl, 2:454–455, 456–457
- Populism, 2:678
- Populist party, 1:251
- Positive liberty, 2:616

- Positive-sum game, 1:407–408
- Positive theory of agency, 1:46
- Positivism and its critique, 2:459–469**
- interpretivism and, 2:467
 - logical, 2:461
 - political science and, 2:453–454
 - scientific legitimacy and, 2:463–467
- Possibility theorem, 1:37
- Postanarchism, 2:629
- Perestroika protest in political science, 1:9–10, 23
- Postbehavioralism, 1:3–4, 8–10, 63
- Postmaterialism, 1:206
- Postmodern constructivism, 2:472–473
- Postmodernism, 1:11–21**
- Baudrillard on, 1:16, 19
 - constructivism and, 1:12
 - cultural shift toward, 1:206
 - Derrida on, 1:14–15
 - Foucault on, 1:15–16
 - future research issues, 1:20
 - Giddens on, 1:12, 13, 16–17, 19
 - Liotard on, 1:12, 13, 14, 19
 - objectivism and, 1:12
 - policy implications of, 1:19–20
 - on social power and oppression, 1:17–19
 - theory on, 1:12–14
- Posttotalitarian state, 1:95
- Poverty, 1:30
- Power. *See* **Balance of power**
- Predominant party system, 1:155
- Prejudice, defining, 2:832
- Presidency, 2:725–34**
- cabinet staffing, 2:727–728
 - campaigns, 2:798
 - Congress and, 2:729–730
 - courts and, 2:731
 - economic policy, 2:732
 - executive branch and, 2:730–731
 - extraconstitutional powers of, 2:726–727
 - foreign policy, 2:733
 - gender and, 2:844–845
 - iron triangle exclusion of, 2:730–731
 - neoinstitutionalism on, 2:727
 - policy implications, 2:731–733
 - psychobiology on, 2:727
 - signing statement threat, 2:733
 - social policy, 2:731–732
 - Straussians on, 1:64
 - transition to new president, 2:727
 - veto threat, 2:730
 - White House staffing, 2:728–729
- Presidentialism versus parliamentarism, 1:177–185**
- cabinet dominance, 1:180
 - democratic instability and presidentialism, 1:182–183
 - endogenously timed elections, 1:82, 181
 - Formateur party, 1:180
 - hybrid regimes, 1:180
 - multiple-party system, 1:181
 - parliamentarism overview, 1:181–182
 - parliamentary regime components, 1:180–182
 - presidential regime components, 1:177–180
 - presidentialism overview, 1:179–180
 - proportional representation, 1:181
 - single-member district, 1:181
 - Westminster model, 1:180, 181
- Presidents, list of U.S., 2:726
- See also* **Presidency**
- Primary sampling unit (PSU), 2:517
- Priming effects, 2:712, 713
- Primitive communism, 2:651
- Primitivism, 2:628
- Primordialism, 1:142, 218, 2:635
- Principal–agent theory, 1:43–50**
- adverse selection in, 1:43, 44
 - application to toxic mortgage debt, 1:48–49
 - classic views of, 1:45–47
 - critiques of, 1:48
 - moral hazard in, 1:43, 1:44–45
 - personal housing market and, 1:44
 - in political science, 1:47–48
 - positive theory of agency in, 1:46
 - tout court, 1:46
- Prisoner's-dilemma game, 1:290, 2:543–544
- Pritchett, C. H., 2:737
- Private property, 1:320
- Probabilistic causation, 1:288, 2:535
- Probability cluster sampling, 2:517
- Probability sampling, 2:519–520
- Probit analysis, 1:289
- Procedural democracy, 1:7
- Procedural rationality, 1:36, 2:816–817
- Professional revolutionary, 2:653–654
- Proletariat, 2:652, 656–657, 665
- Prometheans, 1:443
- Propensity scoring, 2:520
- Proper Sphere of Government, The* (Spencer), 2:596–597
- Proportional representation (PR), 1:160, 1:181
- Proposition, scientific, 2:452
- Prospect theory, 2:535–536
- Protectionist trade policy, 1:409
- Protest cycle, 1:230

Protestant Ethic and the Spirit of Capitalism, The
(Weber), 1:211

Protestant Reformation, 2:578, 583–584, 587–588, 589

Protestant Revolution, 1:210

Protestantism. *See* **Christian political thought**

Proudhon, Pierre-Joseph, 2:626, 649, 655

Przeworski, A., 1:86

Pseudo-science, 2:454–455

Psychoanalytic theory, 1:54, 355–356

Psychobiography, 1:54

Psychobiology, on presidency, 2:727

Psychocultural interpretation, of ethnic identity, 1:142

Psychological effects, in elections, 1:162

Psychological theory of opinion formation and voting,
2:817–818

Psycho-political construction, 2:634

Public administration studies on urban politics, 2:706

Public assistance, 2:597–598

Public choice theory, 1:39, 441

Public education, 1:174, 2:599

Public goods, 1:108

international, 1:425

Public opinion, 2:827–829

Public policy and administration, 2:786–795

agenda setting, 2:788

alternative research models, 2:793

decision-making models of, 2:790–791

evaluation of, 2:792–793

foundation for studying, 2:786–787

future research issues, 2:793–794

implementation, 2:791–792

open-system administration, 2:787

organizational process model of, 2:790–791

policy formulation and adoption, 2:788–790

rational-comprehensive model of, 2:790

stages models of public policy, 2:787–788

Punctuated-equilibrium, 2:793

Pursuit of happiness, 2:688–689

Putin, Vladimir, 1:263

Putnam, R. D., 1:195–196, 197, 199, 202–203,
341, 2:714, 818

Pye, L. W., 2:565

Qualitative content analysis, 2:492

Qualitative research, 1:285, 286, 287, 288, 289

See also **Qualitative versus quantitative research**

Qualitative versus quantitative research, 2:506–513

assumptions of qualitative research, 2:507

assumptions of quantitative research, 2:507

comparing/contrasting, 2:507–510

data analysis in, 2:509

data collection in, 2:508–509

definition of qualitative research, 2:507

definition of quantitative research, 2:507

future research issues, 2:511–512

limitations of qualitative methods, 2:511

limitations of quantitative methods, 2:510

qualitative methods, 2:510–511

quantitative and qualitative research, 2:506–507

quantitative methods, 2:510

reporting of results in, 2:509

research design in, 2:508

research question in, 2:507–508

sampling in, 2:508

summary, 2:509–510

theorizing in, 2:508

Qur'an, 2:570

Quantitative content analysis, 2:492

Quantitative research, 1:285, 286, 287–288, 289–290

formal theory and, 2:533–534

See also **Qualitative versus quantitative research**

Quasi-experimental design, 2:527, 819

Quebecois, 1:143

Queer approach, 2:862–863

See also **LGBT issues and the queer approach**

Rabin, Yitzhak, 1:340

Rabinowitz, G., 2:815

Race, ethnicity, and politics, 2:831–839

accommodationism, 2:833–834

acculturation and, 2:833

African Americans, 2:836–837

assimilation and, 2:833

basic terms and concepts, 2:831–835

deracialization strategy, 2:838

discrimination, 2:832

ethnicity, defining, 2:832

Irish Americans, 2:835

majority and minority status, 2:831–833

Mexican Americans, 2:836

minority political strategies, 2:833–835

prejudice, defining, 2:832

race, defining, 2:832

racial/ethnic political theory, 2:831

racism, defining, 2:832

radicalism, 2:834–835

segregation, 2:832–833

separatism, 2:834

social stratification, 2:832

urban politics and, 2:705

Radical comprehensive approach to bureaucracy, 2:747

Radicalism, 2:834–835, 837

Radicalized modernity, 1:17, 19

Rand, Ayn, 2:618

- Random assignment, 2:524
- Random sampling, 2:517–158
- Rational actor model (RAM), 2:826, 827
- Rational actor theory, 1:38–39
- Rational choice
- congresspersons and, 2:718
 - formal theory and, 2:534–536
 - institutional, 2:793
 - nonstate actors and, 1:419
 - voting behavior and, 2:815–816
- See also* **Rationality and rational choice**
- Rational choice in game theory, 1:285
- Rational choice institutionalism, 1:25, 162
- Rational choice-driven game theory, 1:285, 286
- Rational design and delegation, 1:428–429
- Rational deterrence theory (RDT), 1:369–370, 371–372
- Rational terrorist, 1:120
- Rational-comprehensive model, 1:34, 2:790
- Rationality and rational choice, 1:34–42**
- bounded rationality, 1:35–37
 - contemporary rational choice, 1:38–39
 - debate about, 1:39–40
 - economic rationality, 1:35–36, 39
 - foundations of rational choice, 1:37–38
 - legal rationality, 1:35
 - limits of, 1:35–36
 - objective rationality, 1:35
 - perspective theory and, 1:40
 - policy making/decision making and, 1:34–35
 - political rationality, 1:35
 - in political science, 1:38–40
 - procedural rationality, 1:36, 2:816–817
 - rational actor theory, 1:38–39
 - rational-comprehensive model, 1:34
 - religion and rationality, 1:214
 - social rationality, 1:35
 - Straussians on rational choice, 1:63
 - subjective rationality, 1:35
 - successive limited comparisons model, 1:34–35
- See also* **Principal-agent theory**
- Rawls, John, 2:612, 620, 623
- RDD poll, 2:517–518
- Reagan, Ronald, 2:706, 731, 774, 825
- Realism and neorealism, 1:311–318, 322**
- balance of interests thesis in, 1:315
 - balance of power thesis in, 1:315
 - balance of threat thesis in, 1:315
 - challenges to (neo)realism, 1:315–317
 - classical realism vs. neorealism, 1:313–315
 - criticism of international relations, 1:304
 - economic interdependence and, 1:404
 - international organizations/regimes and, 1:426–427
 - international relations and, 1:304
 - neorealism on distribution of power, 1:37–39
 - neorealism overview, 1:312–313
 - realism overview, 1:311–312
- Realist counterargument, to democratic peace, 1:390
- Redistribution concept, 2:616
- Reflections on the Revolution in France* (Burke), 2:593–594
- Reformation, 2:578, 2:587–588, 583–584
- Reformist fundamentalism, 2:574
- Regime theory
- on nonstate actors, 1:416
 - Straussians and, 1:62–63
- Regional/ethnic parties, 1:151
- Regionalism, impact on trade, 1:411
- Regression analysis, 2:478–89**
- bivariate regression, 2:478–481
 - coefficient of determination (R^2), 2:481–483
 - dummy variables, 2:486–487
 - Gauss-Markov assumptions, 2:484–486
 - interaction terms, 2:487–488
 - multivariate regressions, 2:484
 - ordinary least squares (OLS) regression, 2:478
 - scatterplots in, 2:479–481
 - standard errors of coefficients, 2:483–484
- Rehn-Meidner model, 2:661
- Reification, fallacy of, 1:78
- Religion and comparative politics, 1:209–216**
- ethnicity and, 1:214–215
 - in Europe, 1:212–213
 - evolution of religion, 1:210–211
 - institutional perspective on, 1:215
 - in Latin America, 1:213
 - liberation theology, 1:213
 - in Middle East, 1:214
 - rentier states, 1:214
 - sacred canopy model, 1:212–213
 - scope of, 1:211–212
 - secularization theory, 1:211
 - social capital and, 1:213–214
 - in United States, 1:213–214
- Religion and politics in America, 2:847–855**
- accommodationists, 2:852–853
 - atheists, 2:850
 - evangelicals, 2:853
 - evidence for differences from rest of world, 2:847–848
 - federalism and, 2:853–854
 - importance of religion, 2:848–849
 - market theory on, 2:850–852
 - nondenominational churches, 2:850

- secularization theory on, 2:859
 separationists, 2:852
 supply-side model on, 2:849–850
 Religious law system, 1:187
 Religious organizations, as nonstate actors, 1:418
 Religious sects, 1:143
 Renner, Karl, 2:634
 Rentier effect, 1:86
 See also **Resources and rentierism, global politics of**
 Representative democracy, 1:269–270
 Repression effect, 1:86
 Republic (Plato), 2:557, 558–559
 Republicanism, 1:269–270
 classical, 2:687
 Republicans, 2:772–775
 Reputation theory, 1:373–374
 Resource allocation, 1:72
 Resource dependence conflict, 1:135
 Resource mobilization theory, 1:226–227, 229
Resource scarcity and political conflict, 1:133–140
 as not inevitable, 1:137–138
 climate change, 1:135–137
 measurement issues, 1:137
 resource dependence conflict, 1:135
 trade theoretic model of, 1:134–135
Resources and rentierism, global politics of, 1:393–400
 applications/empirical evidence, 1:395–396
 causal mechanism of rentier behavior, 1:394–395
 foreign aid as rent, 1:396–397
 policy implications of, 1:397–399
 Resurgent ethnic identity, 1:219
 Reverse causality, 2:810
Revisionism and social democracy, 2:655–663
 achievements of, 2:662
 capitalism reform and, 2:660–661
 criticisms of, 2:662–663
 defining social democracy, 2:659–662
 democratic revisionism, 2:656–658
 economism and, 2:662–663
 environmental movement and, 2:662
 expanding democracy by, 2:661–662
 Germany and social democracy, 2:655–658, 660, 661
 move from revisionism to social democracy, 2:658–659
 practitioners of, 2:657
 Sweden and social democracy, 2:658–659, 661
 welfare state and, 2:659–660
 women's movement and, 2:661–662
 Revolution
 continuous, 2:670
 from above, 1:253
 Revolutionary socialists, 2:654
 Revolutionary syndicalism, 2:641, 667
 Revolutionary Armed Forces of Columbia (FARC), 1:110
 Rhodes, R. A., 1:24, 25
 Ricardo, David, 1:407–408, 2:650, 664
 Ricardo–Viner theory, 1:408
 Rice, Condoleezza, 1:382
 Rich, Adrienne, 2:857
Rights of Man (Paine), 2:593, 619
 Riker, William, 1:38, 40, 169–170, 2:609–610
Rivalry, conflict, and interstate war, 1:376–383
 China, policy implications in, 1:382
 cold war policy implications, 1:381–382
 empirical evidence, 1:381
 expected utility theory, 1:379–380
 future research issues, 1:382–383
 Iraq, policy implications in, 1:382
 policy implications, 1:381–382
 rivalry, 1:376–379
 war initiation, 1:379–381
Roe v. Wade, 2:693
 Roman Catholicism. *See* **Christian political thought**
 Romania, fascist movement in, 2:646
 Roosevelt, Franklin D., 2:726, 727, 728, 761, 774
 Roosevelt, Theodore, 2:726
 Rose, Peter, 2:832
 Ross, M. L., 1:134, 398
 Rostow, W. W., 1:101
 Rousseau, Jean-Jacques, 2:592–593, 634, 664, 686
 Rule of law, 1:268
 Runyan, W. M., 1:345, 356
 Russification, 2:668
 Rwanda, 1:433, 2:637
 Sacred canopy, 1:212–213
 Sadat, Anwar, 1:354
 Saint-Simon, Henri, 2:462, 664
 Saltsjobadan Agreement, 2:659
 Salutary neglect, 2:687–688
 Sambanis, N., 1:109, 111
 Satisficing, 1:36, 37, 52
 Saudi Arabia
 rentier effects in, 1:398–399
 See also **Islamic political thought; Middle East**
 Scapegoating, 1:145
 Scatterplots, 2:479–481
 Schema theory, 1:54
 Schlesinger, A., 2:799–800
Science, evolution of in political science, 2:451–458
 behaviorism, 2:453–454
 defining political science, 2:451–452
 falsificationism and deduction, 2:454–455
 interpretive political science, 2:455–456

- positivism and induction, **2:453–454**
- scientific methodologies, **2:452–457**
- synergy, **2:456–457**
- Scientific analysis, **1:77**
- Scientific management theory, **2:747**
- Scope of research, **1:285–286**
- Scott, James, **2:825, 826**
- Secessionist conflicts, **1:109**
- Second Continental Congress, **2:688, 689**
- Secularization theory, **1:211, 2:859**
- Security dilemma, anarchism and, **1:378**
- Seesaw coup, **1:130**
- Segal, J. A., **2:739–740, 741**
- Segmented multipartism system, **1:155**
- Segregation
 - defining, **2:832–833**
 - See also* **Race, ethnicity, and politics**
- Selectorate theory, **1:87, 88**
- Self-adjusting principle, **2:598, 600**
- Self-coup, **1:126**
- Selznick, P., **1:12, 2:791**
- Semiauthoritarianism, 1:258–266**
 - case studies on, **1:263–264**
 - competitive authoritarianism and, **1:261, 262**
 - democracy/autocracy separation, **1:259–260**
 - diminished subtypes in, **1:260, 262**
 - electoral authoritarianism and, **1:260–261, 262**
 - future research issues, **1:264**
 - gray zone in, **1:259–262**
 - gray-zone problem, strategies to solve, **1:260–262**
 - hybrid regimes, quantitative evidence on, **1:262–263**
 - hybrid regimes in, **1:261–262**
 - implications for democracy assistance, **1:264–265**
 - institutionalized semiauthoritarianism, **1:263–264**
 - origins of, **1:259**
 - output legitimacy in, **1:261**
- Semiotics, **1:16**
- Semi-periphery, **1:103, 331**
- Sensitivity analysis, **1:32**
- Separation of powers, **2:692–693**
- Separationists, religion and, **2:852**
- Separatism, **2:834**
- September 11 attacks on U.S., **1:118–119, 120**
- Serbia, **1:145, 2:637**
- Shapiro, I., **1:40**
- Shari'a, **2:571**
- Shaw, Albert, **1:4**
- Shi'a school, **2:570–571**
- Shugart, Matthew, **1:178**
- Shumpeter, Joseph, **1:267**
- Signing statement threat, **2:733**
- Simon, Herbert, **1:35–36, 52, 2:745, 747**
- Simularca/simulations, **1:16**
- Simulation, **1:32**
- Single transferrable vote, **1:160**
- Single-member district (SMD), **1:160, 161, 163, 181**
- Single-party regime, **1:250–251**
- Situational perspective, on ethnic identity, **1:218, 219**
- Six Books of the Commonwealth* (Bodin), **2:585**
- Skinheads, **2:646**
- Skinner, B. F., **1:55**
- Skocpol, Theda, **1:26, 91, 92, 231, 252–253, 2:460**
- Skyjacking, **1:118**
- Slave society, **2:651**
- Slavery, **2:692, 2:836**
- Small-*N* studies, **1:286, 287, 288**
- Smith, Adam, **1:30, 320, 407, 2:618, 650, 664**
- Smith, Rogers, **1:23**
- Social capital, **1:20, 193, 194–195, 202–203, 213–214, 298, 2:818**
- Social Contract, The* (Rousseau), **2:592–593, 634**
- Social contract theory, **2:589–593, 590**
- Social control, **1:15–16**
- Social Darwinism, **2:596, 600**
- Social Democratic Party (SDP), **2:600, 2:655–656**
- Social democracy. *See* **Revisionism and social democracy**
- Social ecology, **2:627–628**
- Social equality, **1:7**
- Social factors, in ethnic conflict, **1:145**
- Social mobilization, modernization and, **1:86**
- Social movements, 1:225–232**
 - civil rights movement, **1:227**
 - cultural research approaches, **1:227**
 - development of social movement studies, **1:226–228**
 - early research approaches, **1:226**
 - framing approach to, **1:228**
 - future research issues, **1:230–231**
 - global, **1:404**
 - green movement, **1:227**
 - grievances approach to, **1:228**
 - participation in, **1:228, 229, 231**
 - policy effects of, **1:229–230**
 - political opportunity structure approach to, **1:227–228**
 - resource mobilization theory on, **1:226–227**
 - strain theory on, **1:227**
 - transnational, **1:229**
 - women's movements, **1:230**
- Social networks, **1:193, 195–196**
- Social policy, presidency and, **2:731–732**
- Social power and oppression, construction of, **1:17–19**
- Social rationality, **1:35**
- Social science, systemism approach to, **1:30**

- Social stratification, defining, **2:832**
 Socialdemokratiska Arbetarepartiet (SAP), **2: 658–659**
 Socialism
 market, **2:678**
 See also **Fascism and national socialism; Socialism in the developing world**
Socialism in the developing world, 2:673–681
 anti-imperialism and anti-exploitation, **2:674**
 barriers to socialist transition, **2:674–675**
 case studies, **2:676–679**
 China, **2:677–678**
 Cuba, **2:676–677**
 economic development and, **2:673–674**
 Iraq, **2:677**
 Libya, **2:678–679**
 North Korea, **2:679**
 problems of variants of socialism in, **2:679–680**
 reasons to adopt socialism, **2:673–674**
 revisionist variant of socialism, **2:677**
 Soviet influence on, **2:674**
 traditional societies, **2:674**
 variants of socialism, **2:675–676**
 Yugoslavia, **2:678**
 Socialist parties, **1:151–252**
 Socialization process, **1:427**
 Society-centeredness approach, **1:91**
 Sociological institutionalism, **1:26**
 Sociological liberalism, **1:323–324**
 Sociological theory of opinion formation, **2:813–814**
Sociology, history of discipline of, 1:3–10
 behavioralism, **1:3–4, 1:6–8**
 Goodnow and, **1:5**
 Perestroika protest, **1:9–10**
 postbehavioralism, **1:3–4, 8–10**
 traditionalism, **1:3–5**
 W. Wilson and, **1:5**
 Socrates, **1:62–63, 65, 2:554, 555**
 Soft balancing (of power), **1:366**
 Soft law, **1:434**
 Somalia, **1:433**
 Sorel, Georges, **2:641**
 Soul, concept in ancient world, **2:555–256**
 South Africa, high courts in, **1:190**
 South Vietnam, **2:674**
 Sovereignty, state
 democratic peace and, **1:384–386**
 globalization and, **1:405**
 Soviet Union
 cold war and, **1:381–382, 2:637**
 dissolution of, **2:612, 637**
 elections in, **1:161**
 influence on socialism in the developing world, **2:674**
 Spain, language issues in, **1:174**
 Spanish Civil War, **2:627**
 Spatial modeling. *See* **Formal theory and spatial modeling**
 Spatial theory of voting, **2:537–538, 818–819**
 Speath, H. J., **2:739, 740**
 Spencer, Herbert, **2:596–600**
 Stages models of public policy, **2:787–788**
 Stalin, Joseph, **2:633–634**
 Stalinism, **2:668–669**
 Standard errors of coefficients, **2:483–484**
 Standpoint theory, **1:19**
 Stark, R., **2:850**
 State
 international law and, **1:435–436**
 will of, **1:5**
 See also **Statism**
State and local government, 2:779–785
 bureaucracy of state government, **2:783**
 county government, **2:783–784**
 governors, **2:781–782**
 municipal government, **2:784**
 state constitutions, **2:780–781**
 state judicial systems, **2:782–783**
 state legislatures, **2:782**
 state political structures, **2:780–783**
 states in federal system, **2:779–780**
 variety and inequality in, **2:780**
 State terror, **1:116**
 State-level campaigns, **2:798**
 State-nation, **2:634, 635**
 States-rights amendment, **2:780**
 Stationary subgame perfect Nash equilibrium, **2:546–547**
Statism, 1:91–98
 authoritarian states, types of, **1:94–95**
 autonomy and, **1:92**
 “bringing state back in” concept, **1:91, 92–93**
 definition of modern state, **1:92**
 domination of state, **1:92**
 nation-state and, **1:93–94**
 politicians’ dilemma and, **1:95–97**
 state building and structure, **1:93–95**
 third-world countries and, **1:94**
 weak state syndrome and, **1:95–97**
 Statistical models, **1:286**
 Steffens, Lincoln, **2:701–702**
 Stephanus numbers, **2:554**
 Stiglitz, Joseph, **1:135, 1:137**
 Stirner, Max, **2:627**
 Stolper–Samuelson theory, **1:408**
 Stone, Clarence, **2:700**

- Stone, Deborah, 2:793
- Storing, Herbert, 1:63, 64, 66
- Strain theory, 1:227
- Strategic choice institutionalism, 1:25
- Strategic model of judicial behavior, 2:736
- Strategic terrorist, 1:120
- Strauss, Leo, 1:60, 2:558
- See also* **Straussians**
- Straussians, 1:60–68**
- classical political science preference of, 1:60–63
- common sense problem and, 1:61–62
- criticism of, 1:65–66
- East Coast Straussians, 1:66
- Midwest Straussians, 1:66
- regime analysis and, 1:62–63
- Straussian studies, 1:63–65
- theological-political problem and, 1:66–67
- West Coast Straussians, 1:66
- Street fighting pluralism, 2:700
- Structural functionalism. *See* **Systems theory and structural functionalism**
- Structural realism, 1:313
- Structural-functional requisite analysis, 1:78
- Structuralism
- dependency theory and, 1:103–104
- historical, 1:409–411
- international trade and, 1:409–411
- on perpetual peace, 1:387
- Structuration theory, 1:17
- Structure, defining, 1:77–78
- Subject political culture, 1:201–202, 2:756
- Subjective rationality, 1:35
- Subtle media effects, 2:712–713
- Successive limited comparisons model, 1:34–35
- Suffrage, 1:153
- Suicide attacks, by terrorists, 1:122–123
- Summa Theologica* (Aquinas), 2:581–582
- Sumner, William Graham, 2:600–603
- Sun Yat-sen, 2:564
- Sunni school, 2:570
- Supply scarcity, 1:243
- Supply-side model, 2:849–850
- Supreme Court, U.S., 1:188, 2:731
- Surplus value, 1:102, 2:652
- Survey questionnaire design, 2:515–516
- Survey research, 2:514–522**
- attitude questions, 2:516
- basic ideas in survey research, 2:514–518
- belief sampling, 2:516
- cellular phones and, 2:518–519
- close-ended questions, 2:515–516
- cross-national surveying, 2:520–521
- face-to-face interviews, 2:517
- factual recall, 2:516
- Mitofsky–Waksberg method, 2:518
- open-ended questions, 2:515–516
- order and wording of questions, 2:516
- primary sampling units, 2:517
- probability and nonprobability sampling, 2:519–520
- probability cluster sampling, 2:517
- propensity scoring, 2:520
- random sampling, 2:517–518
- randomized experiments vs., 2:525–526
- RDD poll, 2:517–518
- survey questionnaire design, 2:515–516
- telephone surveys, 2:517–518
- in urban politics, 1:31
- web survey methodology, 2:519
- Survey-based experiments, 2:529
- Survival analysis, 2:502–503
- Sweden, social democracy in, 2:658–659, 661
- Switzerland, regionally concentrated ethnic groups in, 1:146
- Synergy, 2:456–457
- Syria, 1:174, 2:674
- Systematic policy analysis, 2:792
- Systemism, 1:29–33**
- applications, 1:30–32
- evolution of, 1:29–30
- holism and, 1:29–30
- in social science, 1:30
- in urban politics, 1:30–31
- individualism and, 1:29
- policy and, 1:32–33
- postulates of, 1:30
- system analysis, 1:31–32
- Systems. *See* **Comparative political parties: Systems and organizations**
- Systems analysis, 1:31–32
- analysis and optimization in, 1:32
- implementation in, 1:32
- modeling in, 1:31–32
- problem formulation in, 1:31
- Systems theory and structural functionalism, 1:70–80**
- criticisms of system analysis, 1:74
- functional analysis, applying to politics, 1:75–77
- functional analysis, formal, 1:75
- functional analysis, traditional, 1:75
- functional analysis, types of, 1:75–79
- inputs and outputs in systems theory, 1:73–74, 1:76
- political system, 1:72
- structural functionalism, 1:74–75
- structural functionalism, criticism of, 1:78–79
- structural functionalism, history of, 1:74–75

- system analysis applications, 1:72–74
- systems theory overview, 1:70–74
- terminology, 1:77–78
- Tacit consent, 2:591
- Taft, William H., 2:726, 773
- Tanzania, 1:190
- Tate, C. Neal, 1:186
- Tax system, 1:48
- Taylor, Frederick, 2:747
- Taylor, R. C. R., 1:25, 26
- Teleology, 1:77
- Telephone surveys, 2:517–518
- Temporary Autonomous Zone (TAZ), 2:628–629
- Ten yardsticks of federalism, 1:170
- Tennessee Valley Authority, 2:791
- Term limits, 2:782
- Territorial cleavages, 1:151
- Territorial system, 1:169–170
- Terrorism, 1:115–123**
 - Al Qaeda and September 11 attacks, 1:118–119
 - anarchism as, 1:117
 - as distinct form of political violence, 1:115–117
 - defining, 1:308–309
 - domestic, 1:117–118
 - economic effects of, 1:122
 - effects of, 1:121–123
 - historical examples of, 1:117–119
 - historical patterns, 1:119
 - historical use of terrorism, 1:115–116
 - individual and group motivations in, 1:119–121
 - LGBT issues and, 2:863
 - poetic, 2:628
 - rational terrorist, 1:120
 - September 11 attacks, 1:119
 - skyjacking as, 1:118
 - strategic terrorist, 1:120
 - structural causes of, 1:121
 - suicide attacks, 1:122–123
 - See also* **Anarchism**
- Tertiary associations, in civil society, 1:199
- Teutonic Order, 2:641
- Tewksbury, D., 2:714
- Thatcher, Margaret, 2:661
- Theological-political problem, 1:66–67
- Theory assessment, 1:287–289
- Theory generation, 1:284–285
- Theory of Justice* (Rawls), 2:620,
- Theory of revolutionary syndicalism, 2:667
- Thick rationality, 1:39
- Thick social capital, 1:195
- Thin constructivism, 2:472
- Thin rationality, 1:39
- Thin social capital, 1:195
- Think constructivism, 2:472–473
- Third parties, 2:775–776
- Third wave of democratization, 1:255, 275, 276
- Third-world countries
 - state building and, 1:94
 - See also* **Resources and rentierism, global politics of;**
Socialism in the developing world
- Tibetans, in China, 1:143
- Tickner, J. Ann, 1:346–347, 349
- Tilly, Charles, 1:92, 116, 280–281
- Titoism, 2:678
- Tocqueville, Alexis de, 1:63–64, 81, 194, 298, 2:753
- Top-of-the-ticket contests, 2:798
- Toronto Group. *See* **Comparative environmental politics and conflict**
- Totalitarian state, 1:95
- Totalitarianism and authoritarianism, 1:249–257**
 - autodictatorship, 1:250
 - bureaucratic-authoritarian system, 1:94, 253
 - cultural authoritarianism, 1:255
 - democratization and, 1:255–256
 - empirical questions, 1:252–256
 - fear and coercion in, 1:255
 - future research issues, 1:256
 - holism and totalitarianism, 1:30
 - ideology and, 1:255
 - legitimacy of authority, 1:249–250
 - mass spectacle and, 1:255
 - military regime, 1:251
 - monodictatorship, 1:250
 - mysticism and, 2:641
 - neopatrimonialism, 1:250
 - origins of regimes, 1:252–254
 - party regime, 1:250–251
 - performance-based legitimacy and, 1:254
 - personalist regime, 1:251
 - polyarchy, 1:250
 - predatory state, 1:254
 - rule and survival by regimes, 1:254–255
 - semidictatorship, 1:250
 - sultanism, 1:250
 - theoretical and definitional questions, 1:249–252
 - totalitarian regime, 1:251–52
 - See also* **Semiauthoritarianism**
- Tout court principal-agent theory, 1:46
- Trade. *See* **International political economy and trade**
- Trade theoretic model of resource scarcity/conflict, 1:134–135
- Traditional functional analysis, 1:75
- Traditionalism, 1:3–5

- Traditionalist fundamentalism, **2:574**
 Traditionalist political culture, **2:757**
 Transaction costs, **1:46, 48**
 Transcendental-pragmatic method, **2:610**
 Transfer economy, **1:254**
 Transfer of social technology, **1:402**
 Transnational social movements, **1:229**
 Transsexual/transgender movement, **2:859–860**
 Treaties, international, **1:433–434**
 Treaty of Versailles, **2:565, 824**
 Treaty of Westphalia, **1:432**
 Trial courts, **1:190–191**
 Trust, political, **1:203–204**
 Trust networks and categorical equality, **1:279**
 Tullock, Gordon, **1:37–38, 40, 2:536**
 Turbulence theory on nonstate actors, **1:416**
 Turkey, language issues in, **1:174**
Two Treatises on Government (Locke),
 2:591, 617–618, 686
 Two-and-a-half party system, **1:154**
 Two-party system, **2:772–775**
 Two-way interaction model, **2:791–792**
 Tyranny, **1:64**
- Ukraine, political culture in, **1:341**
Unam Sanctam (Marsilio de Padua), **2:583**
 Unconscious anarchists, **2:629**
 Unified party government, **2:727**
 Unilateralism, **1:366**
 Unipolar system, of balance of power, **1:366**
 Unitary system, **1:168, 171–172, 174**
 United Kingdom
 nuclear weapons in, **1:378**
 See also Great Britain
 United Nations (UN), **1:436–437**
 United Nations Framework Convention on Climate
 Change (UNFCCC), **1:445–446**
 United States
 decentralized education in, **1:174**
 dispersed ethnic groups in, **1:146**
 as federal system, **1:169**
 financial crisis in, **1:402**
 terrorist attacks on, **1:118–119, 120**
 See also **Bureaucracy, American; Federalism,**
 American; Foreign policy, American; Judicial
 politics, American; Political parties, American;
 Political systems, founding of American;
 Religion and politics in America
 United States Agency for International Development
 (USAID), **1:415**
 Universal Declaration of Human Rights (UDHR),
 1:434, 436
- Upanishads, **2:563**
Urban politics, 2:699–707
 at-large vs. district elections, **2:704**
 city within federal system, **2:701–792**
 commission form, **2:703–704**
 council-manager form, **2:703**
 elections, **2:704–705**
 home rule, **2:702**
 mayor-council-chief administrative officer form, **2:703**
 national government and, **2:701**
 national-level policy, **2:706**
 political machines, **2:272, 702, 769**
 political structures, **2:702–705**
 public administration studies on, **2:706**
 public policy, **2:706**
 race and ethnicity, **2:705**
 reform movement, **2:702–704**
 research methodology, **2:701**
 state governments, **2:702**
 strong mayor-council form, **2:703**
 systemism approach to, **1:30–31**
 theories of, **2:699–701**
 wealth and poverty, **2:705–706**
- Urdal, H., **1:134, 137**
U.S. Congress, 2:717–724
 agenda setting, **2:721**
 bureaucracy and, **1:48, 2:721–722**
 constituents will and, **2:718–719**
 courts and, **2:721**
 divided government and, **2:720–721**
 electoral dilemma and, **2:719**
 foreign policy and, **2:825–826**
 future research issues, **2:722–723**
 instructed delegate theory on, **2:718**
 interest groups and, **2:718, 722**
 policy implications, **2:720–722**
 presidency and, **2:729–730**
 public perceptions of, **2:719–720**
 theoretical background, **2:717–719**
 U.S. Supreme Court, federalism cases, **2:767**
 Utility maximization, **1:57, 2:534–535**
 Utopian structures, **1:78**
- Vanguard party, **2:666–667**
 Vattel, Emmerich de, **1:320**
 Venezuela
 natural resources and, **1:85**
 rentier effects in, **1:398–399**
 Verba, S., **1:201–202, 285, 2:457, 757**
 Veto coup, **1:126**
 Veto power, of president, **1:178, 2:730**
 Vienna Circle, **2:453**

- Vietnam, adoption of communism in, **2:674**
 Vietnam War, **1:230, 2:828–829**
Vindication of the Rights of Women, A (Wollstonecraft), **2:593**
 Visibility, as LGBT issue, **2:862**
 Voegelin, E., **2:466**
 Voluntary associations, **1:203**
 Vote of confidence, **1:180**
 Vote of no confidence, **1:180**
Voting behavior, 2:813–821
 biological causes of, **2:819**
 collective action, **2:816**
 D term in, **2:816**
 decision making and, **1:38**
 Erfurt Program, **2:656, 661**
 ethnic groups and, **1:220–221**
 informational shortcuts, **2:816**
 media and opinion formation theories, **2:814–815**
 paradoxes, **2:536–537**
 political attitude research and, **1:55–56**
 procedural rationality, **2:816–817**
 psychological theory on, **2:817–818**
 rational choice theory on, **2:815–816, 819**
 recent research on, **2:818–819**
 social capital and voting, **2:818**
 sociological theories of, **2:813–814, 819**
 spatial theory of, **2:537–538**
 spatial theory of voting, **2:818–819**
 types of elections and, **2:819**
 voter efficacy, **1:163**
 Voting Rights Act, **2:837**
- Wagner, Richard, **2:642–643**
 Wallerstein, Immanuel, **1:101, 103, 331, 332**
 Walt, Stephen, **1:40, 378, 382**
 Waltz, Kenneth, **1:314, 315, 377, 379**
 War crimes, **1:116**
 War Powers Act, **2:825**
 Warm neoclassical liberals on, **2:599**
 Warsaw Pact, **1:354**
 Washington, Booker T., **2:836**
 Washington, George, **2:696, 744, 772**
 Weak state syndrome, **1:95–97**
 Weak states
 civil war and, **1:109, 113**
 ethnic conflict and, **1:144**
- Wealth of Nations* (Smith), **2:618**
 Web survey methodology, **2:519**
 Weber, M., **2:490–291, 743**
 Weber, Max, **1:13, 47, 92, 211, 249–250**
 Weiss, Carol, **2:793**
 Welfare state, **1:236–37, 2:659–660**
 Wendt, A., **1:316, 427**
 Westernization, globalization as, **1:404–405**
 Westminster democracy, **1:180, 270–271**
 Westminster parliamentary system, **1:160, 169**
What Social Classes Owe to Each Other
 (Sumner), **2:600**
 White flight, **2:705**
 Wildavsky, Aaron, **1:35, 2:745, 791, 792**
 Wilson, Woodrow, **1:3, 5, 54–55, 312, 321, 355, 377, 2:773**
 Wolin, Sheldon, **2:466–467**
 Wollstonecraft, Mary, **2:593**
 Women's movements, **1:230, 238–239**
 social democracy and, **2:661–662**
 See also Gender and politics; Gender and politics in the United States
 World Bank, **1:305, 329, 409, 424, 438**
 World Trade Center (WTC), **1:409, 412**
 See also September 11 attacks on U.S.
 World Trade Organization (WTO), **1:305, 334, 407, 412, 424, 438**
 World-systems. *See* **Dependency and world-systems**
 WTB Plan, **2:658**
- Yearbook of International Organizations, The*, **1:424**
 Yeltsin, Boris, **1:178, 263**
 Yoshida Doctrine, **2:565**
 Young, Michael, **1:356**
 Yugoslavia, **2:607–608, 637, 678**
- Zaire, **1:145**
 Zaller, J. R., **2:712–713, 814–715**
 Zerzan, John, **2:628**
 Zhou Enlai, **2:565**
 Zimbabwe, **1:355**
 Zionism, **2:574**
 Zuckert, Catherine, **1:66**
 Zuckert, Michael, **1:66**